Remote Sensing Data for Mapping and Monitoring African Savanna Woodlands

Sizwe Doctor Mabaso

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

October 26, 2016

Aberystwyth University of Wales

Supervisors: Dr. Pete Bunting, Dr. Andrew Hardy, Dr. Sandra Brown and Prof. Richard Lucas

Declaration and Statements

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ..... Date .....

Statement 1

This thesis is the result of my own investigations expect where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in footnote(s). Other sources area acknowledged by footnotes and giving explicit references. A bibliography is appended.

Signed ..... Date .....

Statement 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summery to be made available to outside organisations.

Signed ..... Date .....

Acknowledgements

My heartfelt thanks to my team of supervisors; Pete Bunting, Andrew Hardy, Sandra Brown and Richard Lucas, for every way they inputted and offered support, guidance and encouragement throughout the course of the Ph.D. Special thanks goes to the EO Group for their input into my academic growth, and the kindness and laughter. Nathan Thomas deserves a special mention for his precious time and resources, and especially for the patience in decoding my scripts. The long data analysis and write-up nights and the rabbits on our way out are all in my mind.

I'm grateful to Aberystwyth University for funding the Ph.D. Special thanks is also extended to project partners who ensured the success of the project through data provision and fieldwork support. These include Kongsberg Satellite Services (Norway), Norwegian University of Life Sciences, and Sokoine University of Agriculture.

Elim Aberystwyth was my spiritual home away from home, thanks for making me feel welcome. Thanks especially to the Emeseh family and Joel Pridmore who always looked out for me and Nonto. To the Drive Time team, I appreciate the golf lessons and great sessions I enjoyed with everyone.

Special gratitude goes to my wife who had to be apart from me for the duration of the PhD. The sacrifice is most appreciated Putjununu, I can imagine how much the distance and arrangement took from you. Your strength in and through it all was the reason I kept going. To the whole family and friends who played different supporting and caring roles, thank you.

Above it all, without the gift of life and health I enjoyed while pursuing the Ph.D, I could not have successfully come this far. Praise be to the Lord for the gift of life.

Abstract

Remote sensing data provide unprecedented opportunities for detecting and monitoring forest disturbance and loss. Disturbance and loss have been successfully mapped where cleared land is of sufficient extent to provide discrimination within an image. However, methodologies for mapping and monitoring forest degradation are still lacking, primarily because these features are small in size. In Tanzania, the size and extent of degradation is poorly understood, yet there has been an increase in its drivers, such as shifting cultivation, settlements, logging and charcoal burning of indigenous trees.

This study aimed to establish the extent remote sensing can map forest cover and change for REDD+ monitoring in Tanzanian savanna woodlands. A forest baseline for Liwale in south-eastern Tanzania was derived using the Land Cover Classification System (LCCS) from a 2012 RapidEye image, and validated using LiDAR data from 2012. The baseline and a 2014 RapidEye image were then used to perform an object-based change detection, by first identifying potential change features by automatically thresholding the data using a method based on optimising the skewness and kurtosis for distribution of the class of interest, and then classifying, using the random forests algorithm, the true change features. The change results were validated using 2014 LiDAR data. These methods were then scaled out to courser Landsat imagery.

The study concluded that both high and low resolution optical RS data have great potential for forest monitoring of the Tanzanian savanna woodlands, even though Landsat does not provide the level of detail to accurately depict small-scale change. Components of a remote sensing-based monitoring system are proposed. However, it was noted that for national monitoring, an integration of both high and low resolution data was best. Being a deciduous environment, it was found that seasonality and persistent cloud cover in the region greatly limit the window for appropriate monitoring data. The developed methodology is robust, and thus can be scaled up to a broader national scale and across similar environments in Africa.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Natural Environment and Forests

Land and water are basic but important commodities (Kushwaha and Mukhopadhyay, 2010). An important part of the land resource that has been receiving much attention worldwide is forests. Magdon and Kleinn (2013) state that this is mainly because of the various functions they play, such as their role in carbon cycles, being the basis for livelihoods for rural dwellers, and being a repository for terrestrial biodiversity. They are carbon sinks and sources (Zahabu and Skutsch, 2008), making them pivotal in regional to global carbon cycles (IPCC, 2003; Percy et al., 2003; Moghaddam et al., 2005), as they provide carbon storage and absorption capacity whilst standing. Moreover, they directly provide food and shelter to human beings, livestock and wildlife, and are a crucial source of water, energy, building material and medicinal plants for human beings (Iqbal, 1993; Reynolds, 2006).

1.1.1 Global Scale

Global forests cover over 4 billion hectares (31%) of the total land area. Of these, about 93% are natural forests, whilst the 7% was planted (FAO, 2010b). A global population of about 14 million people directly benefit from forests through formal employment in the forestry sector. Rural populations depend and benefit directly from forests and forest products for their livelihood (Iqbal, 1993). In developing countries, wood-based fuels are the dominant source of energy for more

than 2 billion poor people. A recent study by Köhlin et al. (2012) reported that in the African continent, over 90% of harvested wood is used for energy, and the need for wood fuel in low-income countries remain large for the foreseeable future. Furthermore, there is a lot of reliance on forest resources for building material, which subject those forests to logging activities. However, wood is not the only resource taken from forests; non-wood forest products for health and nutritional needs, and for income are derived from forests by about 80% of the population in that part of the world (FAO, 2010c).

Globally, forest cover area and condition declined throughout recent human history (FAO, 2010b). Even though forests are of great importance, knowledge of their rates of change and resulting greenhouse gas emissions has remained limited. Achard et al. (2002) and Pan et al. (2011) noted that such rates have high uncertainties. In the last three centuries, global forest area is estimated to have reduced by approximately 40%, with 75% of the loss occurring during the last two. This was the norm until 2005 when declines in deforestation started and afforestation increased. During this period of high forest loss, FAO (2010b) reported that about 25 countries had their forests completely disappear, with 29 other countries losing over 90% of forest cover.

The Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) indicates that the forestry sector, mainly through deforestation, accounts for about 17% of global greenhouse emissions. This makes it the second largest source after the energy sector (FAO, 2008). According to the FAO (2010c) report on global forests, between the years 2000 and 2010, deforestation rates were estimated at about 13 million hectares per year. This was dominantly in developing countries,with deforestation, forest degradation, forest fires and slash and burn practices accounting for a majority of carbon dioxide emissions. They also reported that deforestation was mainly taking place in tropical countries, whereas developed countries were experiencing stable or increasing forest areas. Near East and Asian countries initially experienced substantial deforestation, but more recently these forests have been stable.

Between 1990 and 2010, the amount of forestland designated primarily for the conservation of biological diversity increased by 35% indicating a political commitment to conserve forests, and these now account for 12% of the world's forests. Unlike in Europe and North America where forest cover and biomass are currently on the increase, tropical rainforests have continued to experience clearing (Achard et al., 2002), losing about 10 million hectares per annum of land to deforestation. In addition, FAO (2010c) noted that in other parts of the world forests such

as savannas have suffered more from degradation and fragmentation, which has further impaired ecosystem functioning.

Causes and drivers of forest disturbance from deforestation and forest degradation vary from place to place, both regionally and internationally. They include poor forest management practices in production forests, forest fires, overgrazing, over-harvesting of fuel-wood for charcoal production and other non-wood forest products, illegal cutting of timber, forest pest outbreaks and forest diseases (FAO, 2008).

1.1.2 African Forests

In the African continent, woody forests cover approximately two-thirds of the land surface (Skidmore et al., 2010). The continent has the second largest rainforest in the world, after the Amazon basin. Tropical savannas occupy about 20% of the global forested area and 65% of these are in Africa (Shirima et al., 2011). Biggs et al. (2004) reported that all major environmental zones within the Southern African Development Community (SADC) were experiencing varying levels of threat due to climatic change, human population dynamics, and conversion to other uses. Leading drivers of ecosystem change in southern Africa, amongst others, are widespread land use change, and conversion of large areas to cropland and pastures, with accompanying degradation. In the relatively less economically developed areas of Africa/sub-Saharan Africa, contrary to the role forests play in human lives, knowledge of their condition and dynamics across them in response to natural and anthropogenically-induced changes is poorly understood. For example, according to Baccini et al. (2008), even though the African rainforest is the second largest globally, it is the least known in terms of carbon stocks and forest conversion rates.

The region has also seen a rise in desertification, overgrazing being the main cause (FAO, 2008). The united nations environment programme (UNEP, 2005) noted that the climate in the SADC region was highly variable, yet at the same time, seasonal averages were becoming dryer and warmer as a result of long-term climate change. This, in turn, has an effect on the region's socio-economic systems, and especially on water demand. In the long term, this has an effect on irrigation, farming (rain-fed agriculture) and the availability of drinking water. Besides the high rates of forest disturbance in the form of deforestation, the continent has experienced considerable forest disturbance in the form of degradation. The most common forms of degradation in the continent are through activities that result in some form of clearing such as shifting cultivation, selective

logging, and charcoal production. However, Mitchard and Flintrop (2013) noted that evidence suggested that there was widespread woody encroachment in sub-Saharan Africa's woodlands and savannas. Such studies highlight the scale and numerous kinds of forest degradation that are taking place in these forests, and their important consequences for the global carbon cycle and land-climate interactions.

1.1.3 Tanzanian Forests

About 38% of the Tanzania's 886,000 km^2 total land area is covered by forests and woodlands, which provide a habitat for a diverse wildlife (United Republic of Tanzania, 2002; SJP, 2009). They are still the key source of livelihood for most of the rural population which practice crop production, animal rearing and bee-keeping, amongst a range of other activities. Furthermore, they play a major role in the preservation and restoration of biological diversity, water catchments and genetic resources in these unique natural ecosystems (Division of Environment, 2012). Forests and woodlands cover approximately 33.5 million ha, of which about 13 million ha are gazetted as forest reserves. Generally, they have a high monetary value as they contribute a notable percentage to the country's GDP through exports and tourism earnings. In addition, they are pivotal for recycling and fixing of carbon dioxide (CO_2), and conserving important biodiversity. Over 700,000 people are employed in various forest related activities. Their contribution in this area is underestimated as there is considerable unrecorded labour in the collection of woodfuels and other forest related products consumed at household level (United Republic of Tanzania, 2001; Division of Environment, 2012).

Despite their importance and roles played by the forest resources to the country's economy, especially the rural population, there are challenges faced. These notably hamper the sector's development, and they include deforestation, inadequate forestry extension services, inefficiency wood based industries and poor infrastructural facilities, outdated legislation, fragmented administration, lack of stakeholder participation in the management of the resources and resource databases, and outdated or non-existence of management plans for efficient resource use, amongst others (The Government of Tanzania and FAO, 2009; Division of Environment, 2012). There are a lot of other services that different sectors derive from forest resources, and it continues to offer them. These include pasture for livestock, raw materials for industries, preservation of watersheds, source of water for irrigation, generation of electricity, environmental protection, control of soil erosion and

nutrients.

The contribution forests make to human lives, whether directly or indirectly, is underestimated because of unrecorded consumption of wood fuels, bee products (honey and wax), catchment and environmental values and other forest products (Division of Environment, 2012). However, it is estimated that the sector's contribution to the GDP is between 2.3% and 10% of the country's registered exports. This is a figure that is low, taking into consideration that forests and woodlands account for 92% of the total energy consumption in the country, with bio-energy being the main source of fuel for rural population. The performance of the sector is characterized with low capacity utilization despite the country's great forest potential. There are also huge potential for non-wood products such as tourism, game, bee products but these are still unknown and undeveloped. The utilization and management of these resources require multi-purpose forest management (The Government of Tanzania and FAO, 2009).

The lack of sound management leads to unsustainable exploitation of resources, which also leads to their overexploitation or depletion while they are being under-utilized. An aspect that has continued to be a significant unknown in Tanzanian forests has been forest degradation. Even though literature shows that there is degradation occurring in Tanzanian forests, research work undertaken to understand the dynamics and to quantify the degradation occurring in them has been limited. Swetnam et al. (2011) highlighted that knowledge of the carbon storage capacity of many of the lowland bush-type habitats of Tanzania is still very limited. Studies such as Munishi and Shear (2004) and Lewis et al. (2009) were focused on the carbon storage potential of the woodlands and forests, but only laid a foundation for more research that need to be carried out. There is therefore a need to explore the possibility of using space and airborne remote sensing data, which according to Schmidt et al. (2015), provides exceptional opportunities for monitoring and detecting forest disturbances and losses, to advance the techniques that enhance the mapping and detection of forest disturbance (both deforestation and degradation) in Tanzania.

The country, and the SADC region generally need such data for planning and management of its resources. On a broader scale, the data are crucial for early warning, food security, agriculture, disaster prevention and management, forest and rangeland monitoring, environmental planning, watershed catchments management, statistics on natural resources, biodiversity studies, and climate change modelling. These forests are, however, faced with deforestation at a rate of between 130,000 and 500,000 hectares per annum. The leading drivers are agricultural expansion, livestock

grazing, wild fires, over-exploitation and unsustainable utilization of wood resources and other human activities mainly in the general lands. As early as 1995, a Tanzanian country report by Mitawa and Marandu (1995) found that there were tree species that were already threatened and vulnerable for extinction as they were being harvested at a high rate for various uses such as sawn timber, construction timber and carvings. What increased their danger of going into extinction is their low regeneration potential.

Tanzania is categorised as one of the 34 biodiversity hotspots in the world, with about 25% of the land area occupied by crucial ecosystem such as national parks, game and forest reserves, and key wetlands designated as Ramsar sites (BirdLife International, 2002). Even though the actual statistical rate of loss is not documented in Tanzania, according to Mniwasa and Shauri (2001) and the Division of Environment (2012), the country has been experiencing biodiversity loss. Biodiversity loss is amongst the six problem areas identified by the National Environmental Policy, which required urgent attention (The United Republic of Tanzania, 1997). The other five were deforestation, land degradation, deterioration of aquatic systems, lack of accessible, good quality water, and environmental pollution. WRM (2002) noted that biodiversity had been threatened by several direct and underlying processes, with clearing of forest land at a rate of 400,000 hectares per year during the past two decades.

This is a country with a majority of the population still relying on natural resources to sustain their livelihoods. For instance, tourism based largely on its ecosystems, wildlife and landscapes both terrestrial and marine, contribute more than 12% of the national GDP. These earnings accrued from the biodiversity resources are likely to cease with the threats posed by loss of the biodiversity. Expanding agriculture, clearing of forests for charcoal and firewood, climate change, and desertification are the primary causes of loss of biodiversity. Addressing biodiversity loss requires permanent and long term solutions in the form of development and implementation of appropriate policies guidelines, institutional capacity building and deployment of adequate resources to halt and reduce the intensity of biodiversity loss (Division of Environment, 2012).

Tanzania is one of the pilot nations for the Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD+) programme. REDD+ is a policy mechanism that was agreed under the United Nations Framework Convention on Climate Change (UNFCCC), as a collaborative initiative to assist developing countries build capacity for reducing emissions and to participate in a future REDD+ mechanism (UN-REDD, 2011). Emissions are reduced through

changing the use and management of forest lands, which in turn provide co-benefits of biodiversity conservation and livelihood support (Burgess et al., 2010; Danielsen et al., 2011). To be able to meet the requirements for REDD, more research is needed to understand the forests of Tanzania, as well as advance techniques for mapping and monitoring them.

1.2 Problem Statement

Deforestation and forest degradation adversely impact forest biodiversity, the availability of wood and non-wood forest products, soil and water resources, thus often threatening an important safety net for the rural poor as their livelihood is impacted (FAO et al., 2008). Reducing deforestation and forest degradation may play a significant role in climate change mitigation and adaptation. In addition, it holds potential for yielding significant sustainable development benefits. Furthermore, it may generate a new financing stream for sustainable forest management in developing countries (SJP, 2009). According to FAO et al. (2008), achieving cost-efficient carbon benefits through REDD+ would result in the reduction rate of increases in atmospheric CO_2 concentrations. This would in turn effectively buy countries the much needed time to move to lower emissions technologies.

Land use and land cover (LULC) assessment and monitoring of its dynamics are prerequisite for a sustainable management of natural resources, environmental protection, food security, and for providing core data for monitoring and modelling (FAO, 2008). In the wake of the high levels of forest loss resulting from demographic and socio-economic changes that have continued to exert considerate pressure on both forest cover and conditions (Achard et al., 2002), there were renewed efforts to reduce it. More research went into forest mapping and the establishment of forest monitoring systems.

The efforts have also resulted in a shift from the use of traditional methods such as field surveys, map interpretation and ancillary data analysis (as they are time-consuming, date lagged and often very expensive), to the use of earth observation (EO) technology (Xie et al., 2008; Malatesta et al., 2013). Remote sensing technology has the potential of systematically observing a landscape at various scales and times, which makes it well suited for resources monitoring and management. It offers a practical and economical platform for studying vegetation cover changes, especially over large areas, and has been proven very useful for mapping and detecting forest disturbance (Myint

et al., 2011; Dupuy et al., 2012). This has resulted in the EO technology being one of the preferred approaches for vegetation mapping and monitoring, both at a local to global scale.

Deforestation and forest degradation contribute close to 20% of anthropogenic greenhouse gas emissions globally. Deforestation, probably due to its nature and scale of clearing got much attention (Munishi and Shear, 2004), especially in tropical forests as documented in numerous studies such as Achard et al. (2002), Gerwing (2002), Asner (2009), Harris et al. (2012) and Carreiras et al. (2014). Tropical forests have been, to some extent, well researched and documented, and many countries have worked hard to curb deforestation rates (Guyana Forestry Commission, 2012; FAO and JRC, 2012; Carreiras et al., 2014).

Forest disturbance resulting from degradation was initially not given the same priority as deforestation. However, recent literature shows increased efforts towards understanding and documenting forest degradation dynamics and scales. An observed limitation though is that those efforts are still mainly focused on the tropical forests, with examples being Lambin (1999), Baccini et al. (2008), Joseph et al. (2010), Gourlet-Fleury et al. (2013) and Hirschmugl et al. (2014). Other environments such as the African savannas and woodlands that support large rural populations have suffered more from forest degradation than deforestation. In Tanzania, increased logging of indigenous trees, charcoal burning, cultivation, grazing and settlement were found to be the leading drivers of forest disturbance (Mitawa and Marandu, 1995; FAO, 2010c). This poses a challenge as methodologies for mapping and monitoring forest degradation are still lacking (Munishi and Shear, 2004; Lewis et al., 2009; Swetnam et al., 2011), meaning that there is limited documentation and understanding of its scale and dynamics in the country.

One of the main problems with studying degradation is that these features are small in size, which makes them harder to map in an image compared to deforestation features (Böttcher et al., 2009). The true size and extent of degradation in most environments is therefore currently unknown, yet such knowledge is pivotal for both determining and monitoring change in the forests. There is a great need for the scientific world to understand both deforestation and degradation better in order to inform sound and sustainable decision and policy making. It is also essential to put in place robust, transparent, replicable and long-term monitoring systems, and these systems produce reliable data and information on disturbance for forests protection and preservation (Magdon and Kleinn, 2013).

With as high as about 90% of Tanzanian energy being sourced from wood-based fuels (FAO, 2010a),

this underlines the unaccounted for depletion that may be taking place in these forest resources. To effectively reduce and/or eliminate forest disturbance resulting from both deforestation and degradation, it is crucial to understand and advance methods/techniques that will investigate and document its scales and dynamics. This research therefore, sought to advance methodologies for mapping and monitoring forest disturbance in Tanzanian forests that are currently lacking, through the use of remote sensing. The thesis focused on exploring the appropriateness of the spatial and temporal frequencies of the RapidEye sensor, in singularity and for multi-dates, for forest/non-forest mapping, forest disturbance detection and carbon loss estimation. It further sought to explore the possibility of scaling out the developed methodologies to a medium resolution and freely available Landsat optical imagery. The forest baseline and forest change products were validated using Google Earth imagery and the original optical data image. In addition, these products were assessed for comparability to Light Detection And Ranging (LiDAR) (Bater and Coops, 2009; Bunting et al., 2013b; Fayad et al., 2014).

1.3 Aim and Objectives

The primary aim of this research was to develop a method for mapping and monitoring savanna forests, including degradation, using remote sensing data, and in the process explore the appropriateness of remote sensing data for mapping and monitoring Tanzanian forests for REDD+.

The objectives of this thesis were:

1. To develop a method for establishing a forest/non-forest baseline for Tanzanian open woodlands (miombo) from high resolution remote sensing data.
2. To develop a method for detecting forest change and for classifying the potential change features from high resolution remote sensing data.
3. To investigate the potential of lower resolution Landsat imagery for forest monitoring, specifically, concerning what information is lost with the change in resolution for these woodlands.

1.4 Research Questions

In achieving the stated objectives, this study sought to answer the following research questions:

1. To what extent can a baseline for establishing forest/non-forest be established using high resolution remote sensing data?
2. To what extent can change in forest extent be detected using remote sensing data and can this lead to the development of a monitoring system?
3. Can lower resolution Landsat imagery be used for forest monitoring of savanna forests in Tanzania, if so, what information is lost compared to the higher resolution imagery?

1.5 Thesis Structure

Chapter 1 introduces the problem, then outlines the aim and objectives of the research. Chapter 2 is focused on the research background; a review of literature. Within this chapter, the Tanzanian REDD+ status quo was also reviewed in depth. An analysis of current projects, and especially lessons learnt from them were documented. Chapter 3 provides an overview of environments in south-eastern Tanzania, and then focuses on the Liwale study site. Chapter 4 outlines a detailed account of the preprocessing steps undertaken to prepare the remote sensing data used in the research, which was mainly undertaken using the ARCSI and SPDLib software packages. It details how RapidEye and Landsat datasets over Liwale were corrected to surface reflectance, and how LiDAR point data was processed to produce canopy height models.

Chapter 5 focuses on how preprocessed wet season 2012 RapidEye images were used to produce a forest baseline for Liwale. It details how the potential of RapidEye for discriminating between forest and non-forest areas was evaluated. Chapter 6 is an account of how a method that uses high resolution remote sensing data for detecting potential change areas, and then classifying the resultant potential change features into either change or non-change, using a classification-to-image approach, was developed. Chapter 7 documents how the validated 2012 RapidEye baseline and 2014 RapidEye change map were used to train the scaling out of the forest baseline establishment and change detection techniques, respectively, to the freely available medium resolution Landsat data and thus explore to what level medium resolution data can be used for African savanna forest mapping and monitoring. Then Chapter 8 discusses the findings of the research and their implications for Tanzanian REDD+. It also gives recommendations believed would further advance methods for mapping, quantifying and understanding the long-term impacts of disturbance/changes in Tanzanian forest systems. Lastly, Chapter 9 are the major conclusions of the research.

Chapter 2

Literature Review

2.1 Background

The United Republic of Tanzania is a developing country (United Republic of Tanzania, 2002), and had not carried out systematic national forest inventories in the past. This was mainly due to inadequate technical resources and/or limited finances (United Republic of Tanzania, 2012). 2008 saw the birth of the UN-REDD programme, a policy mechanism that was agreed under the UNFCCC, that aims at reducing carbon dioxide emissions from developing countries. It does this through changes in the use and management of forest lands, while providing co-benefits of biodiversity conservation and livelihood support (UNFCCC, 2010). Tanzania being one of the REDD pilot countries, attracted a significant amount of funding (US$93.5 million) for forest conservation and management between 2009 and 2012 (Kaijage and Kuhanwa, 2013).

Natural resources are the most important natural capital for sustainable livelihoods in Tanzania and most other developing African countries (UNEP, 2005; FAO, 2012), especially for rural populations for whom forests are a critical source of livelihood. For example, current statistics revealed that the biomass resource is depleting at a high rate as about 90% of energy consumed is in the form of biomass (FAO, 2010a; Köhlin et al., 2012). This leaves only 10% to alternative, cleaner and more sustainable energy sources in the country. These include imported petroleum products (8%); electricity from hydro-power, natural gas, biomass co-generation, and coal power plants (1.5%); wind, solar and other energy resources (0.5%). This highlights the urgent need to prioritise

environmental protection and sustainable management of land and forest resources for sustained future economic growth, a move that will ensure sustained livelihoods for local communities in the African continent (Division of Environment, 2012).

Tanzanian forests resources are predominantly owned and managed by the central government through the Forestry and Bee-keeping Department (FBD) in the Ministry of Natural Resources and Tourism (MNRT). The ministry is responsible for the management of natural and cultural resources, as well as for the development of the tourism industry (The Ministry of Natural Resources, 2012). The ministry, in partnership and cooperation with other ministries, non-governmental organisations (NGOs) and international partners, has undertaken substantial work with regard to inventories of natural resources, and management systems. For some resources such as minerals, good structures have been put in place with good management, and these resources are being exploited in a controlled and sustainable way, or at least are accounted for. However, there is still more work to be done on the forest resources to put up monitoring and management structures that will ensure their sustained exploitation.

Winrock International and TerraCarbon (2013) concluded that forests contribute significantly to local livelihoods and the Tanzanian national economy. It further reported, however, that such a contribution was largely unrecorded and thus unrecognised. This conclusion highlights the need for more scientific research that would give more understanding of these forests; their state and role on the society. It is in recognition of this situation that most policy interventions, strategies as well as multilateral environmental agreements are focusing their efforts on issues of sustainability. This has birth programs such as REDD+ (UN-REDD, 2011) that have been crucial in the drive towards the sustainable use of global forests and the reduction of emissions resulting from the disturbance of forests.

2.2 REDD+

2.2.1 Introduction

The United Nations Collaborative Programme on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (UN-REDD Programme), is a collaborative initiative that was launched in 2008 to assist developing countries to build capacity to reduce emissions and

to participate in a future REDD+ mechanism (UN-REDD, 2011). It builds on the expertise of three UN agencies, namely: the Food and Agriculture Organization of the United Nations (FAO), the United Nations Development Programme (UNDP) and the United Nations Environment Programme (UNEP). These three also play the convening role for the Programme. It was created in response to the UNFCCC decisions on the Bali Action Plan and REDD at COP-13 (FAO, 2008). Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD+) is a policy mechanism that was agreed under the United Nations Framework Convention on Climate Change (UNFCCC). Specifically, it aims at reducing carbon dioxide emissions from the developing countries through changes in the use and management of forest lands, while providing co-benefits of biodiversity conservation and livelihood support (UNFCCC, 2002; Burgess et al., 2010; Danielsen et al., 2011).

2.2.2 Climate Change and REDD+ Process

After the original formulation of the project in 2007, the 13^{th} Conference of Parties (COP) of the UNFCCC in Bali brought sustainable forestry development to the centre of the international development agenda. It led to high expectations for reducing emissions from deforestation and forest degradation in developing countries in UNFCCC COP negotiations. REDD+ and carbon markets in general, are perceived as having significant potential for increasing funding for the forestry sector in developing countries. However, the perspective of a REDD mechanism is also highlighting the need for more accurate information on forest resources and for appropriate methods for achieving sustainable forest management in these countries. The Bali meeting called for member countries and the international community to demonstrate on REDD+ alternative methodologies (Sundström, 2010; UN-REDD Programme, 2015).

Additionally, REDD-related initiatives, such as UNREDD, the Forest Carbon Partnership Facility (FCPF) or Forest Investment Programme (FIP) by the World Bank, and the voluntary carbon trade market and private financing in REDD will require commonly agreed and reliable monitoring, reporting and verification (MRV) systems. The scope and quality of the required forest resources information remains a major challenge for developing countries that account for a large proportion of the world's most vulnerable forests. National forest monitoring and assessment, and remote sensing studies and reports have become increasingly important information sources for policy and decision makers and managers at national, regional and international levels (Sundström,

2010).

The UN-REDD Programme assists developing countries in preparing and implementing national REDD+ strategies and mechanisms. These efforts help countries develop the capacity to implement REDD+ strategies and become 'REDD-ready'. It is aimed at providing practical experience and lessons learnt that informs the international dialogue on REDD+ mechanism post 2012 (Olsen and Fenhann, 2009).

2.2.3 Carbon Monitoring through the REDD+ Initiative

The UN-REDD Programme has been designed to support nationally-led REDD+ processes, while promoting an informed and meaningful involvement of its stakeholders both at national and international REDD+ implementation level. This is a programme that was tailor-made to ensure that such participation and involvement include the Indigenous Peoples and other forest-dependent communities (UN-REDD, 2011; UN-REDD Programme, 2015). It further works towards building international awareness and consensus about the importance of including REDD+ mechanisms in a future climate change agreement. At national level, the programme supports REDD+ readiness efforts in partner countries in Africa, Asia-Pacific and Latin America. The support may be direct; whereby partners are assisted in the design and implementation of UN-REDD National Programmes, or complementary; whereby partners receive support to their national REDD+ action through common approaches, analyses, methodologies, tools, data and best practices developed through the UN-REDD Global Programme (UN-REDD Programme, 2015).

There are a number of other initiatives working in partnership with the UN-REDD Programme in assisting countries that are engaged or wishing to engage in REDD+ activities. These include Norway's International Climate and Forest Initiative, the World Bank's Forest Carbon Partnership Facility, the Global Environment Facility, Australia's International Forest Carbon Initiative and the Collaborative Partnership on Forests (UN-REDD, 2011). The focus is on national action delivered through these Joint Programmes, and these reinforced by supporting measures at both regional and global levels (FAO, 2008).

2.2.4 National programmes

Significantly reducing emissions from deforestation and degradation requires a strong global partnership to create a REDD+ mechanism under the UNFCCC is key (UN-REDD, 2011). Kissinger et al. (2012) noted that the long-term viability of REDD+ depends on changing the norm in sectors currently driving greenhouse gas (GHG) emissions from forest lands. If the REDD+ mechanism is to be a success at the global scale, with minimal international leakage, Burgess et al. (2010) noted the importance of full participation of the developing countries with a larger proportion of the world's forest. There is a significant potential being offered by the proposed mechanisms for REDD+ for forests conservation with the aim of reducing negative impacts of climate change (Burgess et al., 2010). Initially, there were nine pilot countries for the UN-REDD, namely; Bolivia, Democratic Republic of Congo, Indonesia, Panama, Papua New Guinea, Paraguay, Tanzania, Viet Nam and Zambia. Currently, the number of UN-REDD partner countries stands at, with a further 39 countries not receiving direct support to their national programmes (UN-REDD Programme, 2015).

Most countries reported weak governance and institutions in forest-related sectors, conflicting cross-sectoral policies and illegal activities (related to weak enforcement) as critical underlying drivers of deforestation and degradation in their Readiness Preparation Proposals (R-PPs) (UNFCCC, 2010). A fundamental aspect of R-PPs is understanding the drivers of deforestation and forest degradation, for a development of sound and viable policies, as well as measures aimed at altering current trends in forest activities (Hosonuma et al., 2012). UNFCCC negotiations encouraged countries to identify land use, land use change and forestry activities, with a focus on those linked to these drivers. Moreover, the need to assess their potential contribution to the mitigation of climate change was highlighted (UNFCCC, 2010). Such an understanding is critical as it informs how forests are changing, as well as inform policies, national REDD+ strategies and implementation plans (Boucher et al., 2011; Salvini et al., 2014). Deforestation and forest degradation are also major components of the anthropogenic carbon cycle, an area with high uncertainty with estimates often excluding processes leading to deforestation and degradation (Houghton, 2010).

2.3 REDD+ in Tanzania

2.3.1 Tanzanian Bilateral Agreement with Norway Government

The Republic of Tanzania is one of REDD countries (The Institute of Resource Assessment, 2010), which has availed opportunities for external funding towards research work that advances the understanding of Tanzanian forests; their management and monitoring. One such funding led to the Government of the Republic of Tanzania entering into a bilateral agreement with the Government of Norway on climate and forest. The agreement resulted in projects aimed at improved measurement, reporting and verification (MRV) for Tanzanian forests using space and airborne remote sensing data. Through the Norwegian Space Centre, the project also cooperates with international partners through the Group on Earth Observations Forest Carbon Tracking (GEO-FCT) task. As a component of the bigger project, this research work aimed at exploring the usefulness and appropriateness of remote sensing data for mapping and monitoring Tanzanian forests, with a focus on advancing methods for mapping and monitoring degradation through the use of high and medium resolution remote sensing data.

Currently, there are a number of thematic case studies being used to fill in information gaps in deforestation rate, contribution of forest sector in the national economy including socio-economic data, as well as alternative livelihood activities, and the carbon stocks (United Republic of Tanzania, 2012). The Ministry of Natural Resources and Tourism also recently completed a National Forest Resources Monitoring and Assessment (NAFORMA) of the state of forest resources, which was done in collaboration with FAO, and using support from the Finnish Government (The United Republic of Tanzania, 2010). A crucial aspect of the REDD+ readiness process in Tanzania is the nine REDD+ pilot projects implemented across the country with funding from the Norwegian Embassy. They have produced a wealth of knowledge and experience on how to implement REDD+ projects that needs to be used for the further development of national structures for REDD+ (United Republic of Tanzania, 2009; The Ministry of Natural Resources, 2012).

2.3.2 REDD+ Status and Financial Flows for REDD in Tanzania

Tanzania is one of the nine pilot countries for the United Nations REDD Programme. For establishing REDD+ actions within the country, it receives donor funding from Norwegian, Finnish and

German governments, as well as benefits from participating in the World Bank's Forest Carbon Partnership Facility. The Norwegian government committed about US$ 80 million for supporting the development of a national REDD strategy, sub-national pilot projects, research and capacity building, improvement of measuring, reporting and verification (MRV), as well as community and private sector engagement. It further receives US$ 4.28 million from the UN-REDD, another programme largely funded by Norway. The government of Finland also contributed US$ 5 million for the development of a national forest monitoring system. The German Climate Change Initiative donated US$ 3.5 million for the improvement of forest management in the Eastern Arc Mountains. Initially, the country also received funding from the World Bank's Forest Carbon Partnership Facility for the development of an RPP (Kaijage and Kuhanwa, 2013). Figure 2.1 shows an overview of finances sourced for REDD+ in Tanzania and their flow between donors and recipients between the years 2009 and 2012.

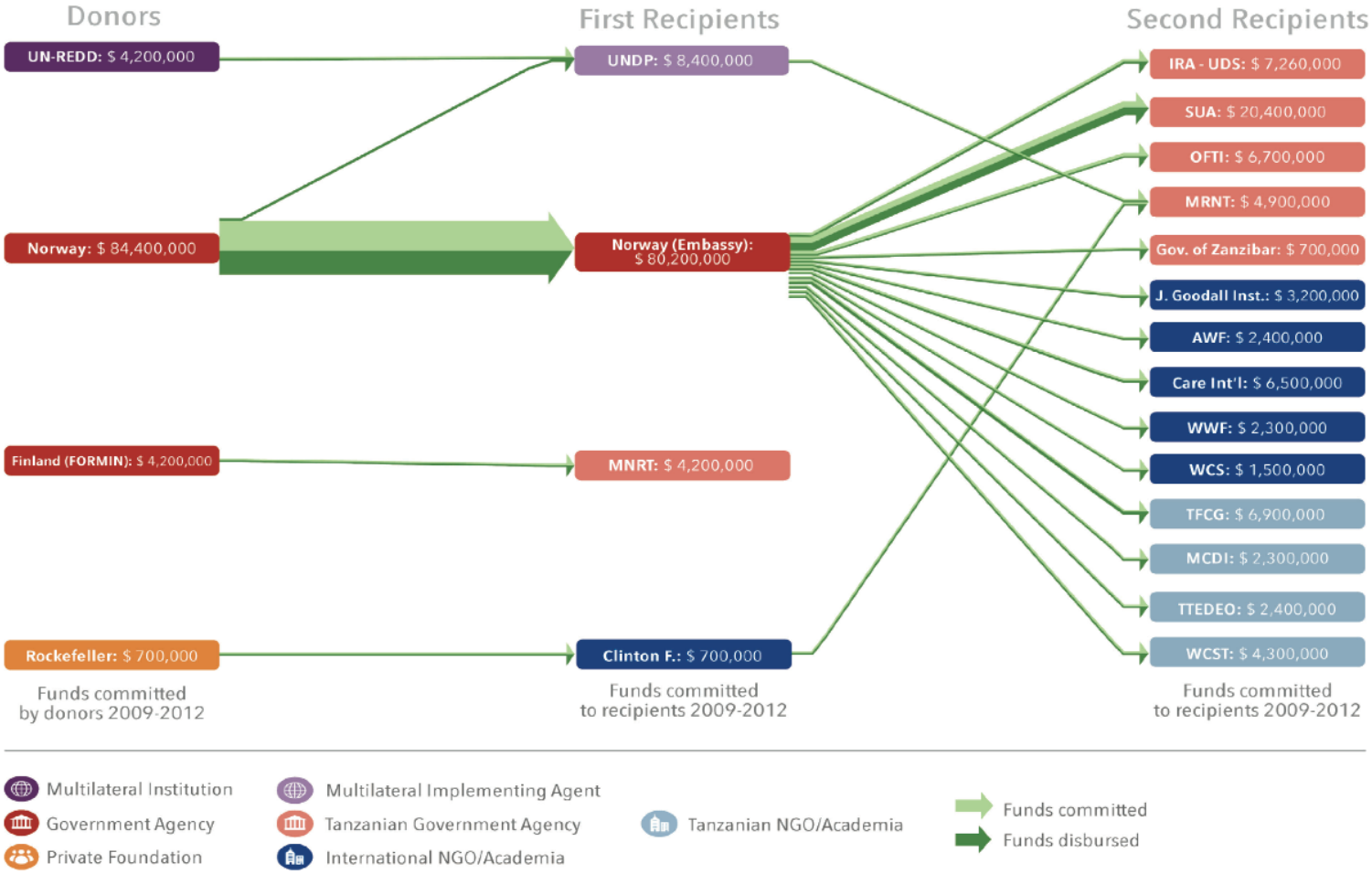


Figure 2.1: An overview of financial flows for REDD+ in Tanzania (2009 - 2012).
Source: Kaijage and Kuhanwa (2013)

2.3.3 On-going REDD+ pilot projects in Tanzania

As part of Tanzania's national REDD+ programme development, there are eight REDD+ pilot projects being implemented by Tanzanian villages and communities, with the aid of civil society organizations. These are all being supported by the Norwegian Embassy in the country (Edwards et al., 2012). These pilot projects cover different regions of Tanzania, and have diverse aims and approaches. There are projects that are being implemented in community-owned forests through Community Based Forest Management (CBFM). There is another set focused on establishing REDD+ in district or central government forests for which adjacent communities have agreed to take some management responsibilities in exchange for forest related benefits (Joint Forest Management). Then there is one project focused on establishing the foundations for REDD+ in protected areas, those being the establishment of carbon baselines, and the carbon monitoring, reporting and verification (MRV). The last project focuses in integrating REDD+ with a customary forest regeneration approach, a process involving small forest plots owned by individuals or institutions within a village, or by the village government itself (Tanzania Natural Resources Forum, 2012). These projects are assisting forest communities to establish the foundations for REDD+, and these include the securing of village land and forest tenure, measuring baselines, establishing of mechanisms for ongoing carbon MRV, and designing benefit sharing mechanisms for anticipated REDD+ revenues. Every project has been found to be unique, and thus with widely varying opportunities and challenges because of the different communities, forest ecosystems, economies, and tenure regimes involved (Ryan et al., 2012; United Republic of Tanzania, 2013b).

2.4 Forest Definition

2.4.1 Working Forest Definition

Common elements in most forest definitions are minimum crown cover, minimum area, minimum tree height, and the absence of other land uses (Magdon et al., 2014). Amongst the most used forest definitions is the one by UNFCCC (2002) that defines a forest as a minimum area of land of 0.05-1.0 ha with tree crown cover of more than 10-30%, and with trees having the potential to reach a minimum height of 2-5 m at maturity in situ. This definition further states that a forest may consist either of closed forest formations, where trees of various storeys and undergrowth cover

a high proportion of the ground, or open forest. Young natural stands and all plantations which have yet to reach a crown density of 10-30% or tree height of 2-5 m are included under forest, as are areas normally forming part of the forest area which are temporarily unstocked as a result of human intervention such as harvesting or natural causes, but are expected to revert to forest (UNFCCC, 2002).

Another commonly used forest definition in literature is one that was adopted by UNEP's ad hoc technical expert group on forest biological diversity (UNEP, 2001). It states that a forest is a land area of more than 0.5 ha, with a tree canopy cover of more than 10%, which is not primarily under agriculture or other specific non-forest land use. This definition further state that in cases of young forest or regions where tree growth is climatically suppressed, those trees should have the capability to reach 5m height in situ, and of meeting the canopy cover requirement (UNEP, 2001).

The third commonly used forest definition is the FAO definition, very similar to that of UNEP, states that a forest is land spanning more than 0.5 ha with trees higher than 5 m and a canopy cover of more than 10%, or trees able to reach these thresholds in situ. This definition excludes land that is predominantly under agriculture or urban use (FAO, 2006b). This is a global definition that FAO has been using to undertake the Global Resources Assessments (FRA) at regular intervals.

For this research work, the working forest definition was adopted from the UNFCCC forest definition, in line with the adopted Tanzanian definition for REDD+. It is defined as a vegetated area with a minimum land area of 0.05 ha, with a tree crown cover of more than 10% and with the potential to reach a minimum height of 2 m at maturity in situ (UNFCCC, 2002). It follows that forests can either be open or closed, depending on the cover of trees and undergrowth, including young natural stands and plantations that have yet to reach the minimal crown cover of 10% or tree height of 2 m. It also comprise of areas that are temporarily disturbed as a result of natural events (e.g. wild fires) and human interventions (e.g., harvesting), but which are expected to revert to a forest.

For space-borne remote sensing studies, this forest definition is a significant challenge to accurate measure and monitoring. For field studies or with the use of very high resolution airborne data, measurements can be taken which meet this definition. However, field base or airborne remote sensing forest monitoring is time consuming and expensive (Næsset and Økland, 2002; Cho et al., 2012; Weng, 2012; Ban et al., 2015). Therefore, in this thesis, it is judged that this forest definition

was extremely challenging and therefore not widely achievable.

A forest is non-degraded if it is still natural, pristine, and has not been disturbed by direct anthropogenic processes, and thus still fully-stocked and functioning naturally (Schoene et al., 2007). However, with increased world populations, and an increased strain on natural resources, previously pristine areas have experienced a rise in both natural and human disturbance. This has usually resulted in a disturbed forest, be it as a result of partial or complete removal of individual trees or groups of trees as a consequence of a discrete event. However, the conversion of a forested land into non-forest land, with a removal of tree cover to less than 10% is classified as deforestation.

2.5 Forest and Disturbance

Numerous studies such as Iqbal (1993) and FAO (2006b) highlight the important role forests play as they directly provide food and shelter to human beings, livestock and wildlife, as well as being a source of water, energy, building material and medicinal plants. For sustainable use of forest resources, there is a need to generate sound information about the forest resources and ecosystems, which holds true for both the scientific community and decision makers; for the operationalisation and implementation of international environmental conventions (Magdon et al., 2014). The development of renewable energy production strategies, globalisation of wood products market, and the need for carbon stock estimates used in climate change modelling has led to the great demand for explicit spatial forest resources data (Lu, 2006; Gallaun et al., 2010).

Hosonuma et al. (2012) reported a huge rise in the need for national data on the type and relative importance of the drivers of both deforestation and degradation. Small scale farming, fuel-wood collection and charcoal production are among the leading drivers of forest disturbance and, ultimately, the reduction of vegetation carbon stocks in African woodlands (Ryan et al., 2012). Forest disturbance leading to forest degradation globally is estimated to affect an area ten times more than the area affected by deforestation per annum (FAO, 2006a, 2008). However, it was noted that the processes that lead to forest disturbance, and thus the reduction of forest stocks have not been adequately characterized spatio-temporally. Information on drivers and activities causing deforestation and degradation is lacking.

Assessment of forest characteristics such as forest structure, aboveground biomass, and forest den-

sity among others, which has been improved especially for management and monitoring purposes (Lee et al., 2001; Chave et al., 2004, 2005; Midgley et al., 2010; Mitchard et al., 2011; Bryan et al., 2013). Quantification of forest parameters in different successional stages, according to Wijaya et al. (2010), helps to inform global emissions and ecosystem changes. A vital aspect of forestry studies entails understanding woody plant dynamics which, according to Midgley et al. (2010), is important from both a theoretical and management perspective. This understanding ensures that informed, complete and unbiased decision-making and management practices are adopted, as well as enhance our understanding of the processes forests go through (Mohammadi and Shataee, 2010).

There are different forest definitions in literature, partly because of many forest types in different environments, but also because different people identify and define them differently (Schoene et al., 2007). Research shows that deforestation is the major disturbance experienced in tropical forests, yet forest degradation has been found to be the major disturbance experienced in savanna forests. Tanzania is one country that is experiencing substantial forest degradation, which unfortunately, is unaccounted for (Division of Environment, 2007; United Republic of Tanzania, 2013a). However, a country report by Kweka et al. (2015), even though not presenting the annual rate of degradation, estimated the annual carbon dioxide emission due to degradation to be 26,447,996 tonnes. Like most developing countries where forest inventories and baselines are missing, the rate and scale of the degradation is poorly accounted for, or missing altogether. The research focused on advancing methodologies for mapping and monitoring Tanzanian forests, a step aimed at improving the measurement, reporting and verification (MRV) methods for these forests. To undertake the task effectively, there had to be clear working definitions of terms. This section focused on conceptualising and giving working definitions for these terms that are a backbone of the; they include forest, forest disturbance, deforestation, and forest degradation.

2.5.1 Deforestation and Forest Degradation

Literature defines deforestation as the long-term or permanent conversion of land from forested to non-forested. FAO refers to forest degradation as changes taking place within the forest, negatively affecting the structure or function of that forest stand (FAO et al., 2008). These changes have to be to the extent that the tree canopy cover goes below the minimum 10% threshold, and thus lower the capacity to supply products and/or services. UNFCCC (2002) defines it as the direct human-

induced conversion of forested land to non-forested land. The FAO defines it as 'the conversion of forest to another land use or the long-term reduction of the tree canopy cover below the minimum 10% threshold' (FAO et al., 2008). It is basically the replacement of forests with different land cover types or uses (Bryan et al., 2013), into less bio-diverse ecosystems such as pasture, cropland, or plantations, urban use, logged area, or wasteland (Ahmed, 2008).

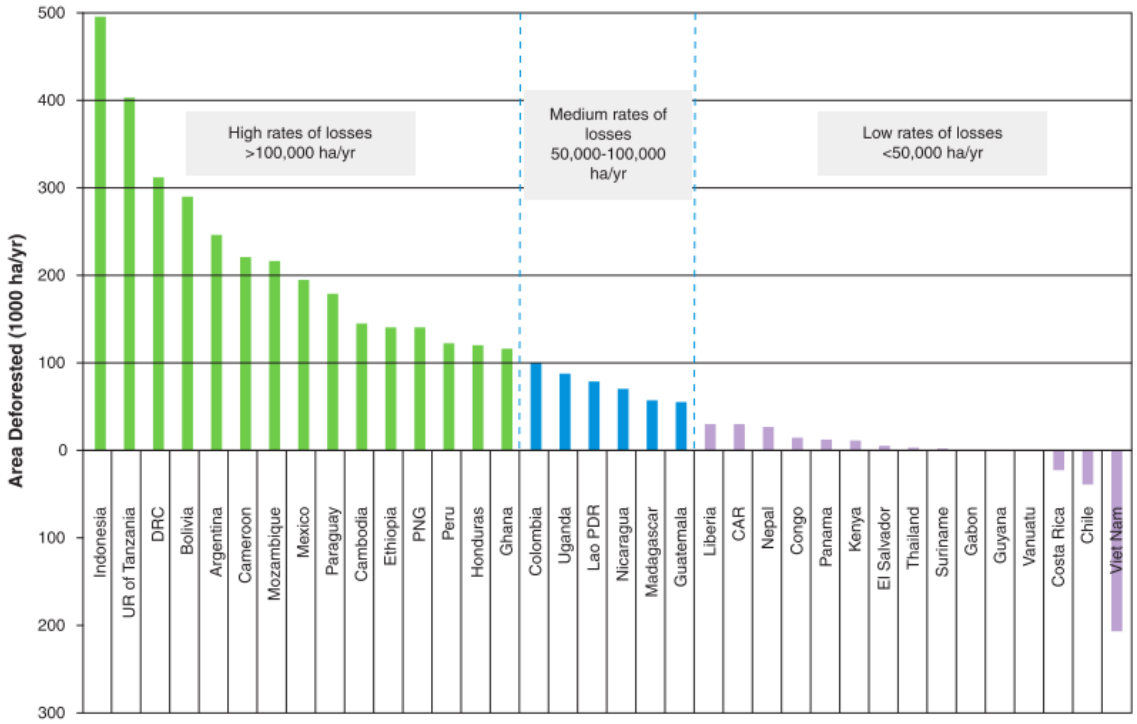


Figure 2.2: The deforestation rates of developing world countries where Tanzania is listed as having the second highest rate.

Source: FAO (2010c)

Unlike deforestation, forest degradation is a process whereby a forest is repeatedly disturbed or undergoes adverse environmental conditions that result in any important forest values being lost or reduced (IPCC, 2003). The modification may be temporary or permanent; in a time-scale of a few years to a few decades. The original forest type is rarely replaced by any other forest (Lambin, 1999). However, this modification is usually associated with reduced forest components, increased emissions of carbon and other greenhouse gases and reductions in the capacity of forests to provide goods that Iqbal (1993) list some as timber, water, plants and animals. Moreover its capacity to offer critical services such as cycling of water, protecting soils, harbouring biodiversity and collectively contributing to biological and ecological functioning is reduced (Lambin, 1999; IPCC,

2003; Herold and Román-Cuesta, 2011). Bryan et al. (2013) defines it as the 'substantial reduction' of biomass, commonly from the big trees, but still retaining sufficient tree cover to remain a forest. Herold and Román-Cuesta (2011) called it the 'loss of carbon stock within forests that remain forests'.

Process and Causes of Forest Disturbance and Degradation

Deforestation and forest degradation are closely related as both can cause each other. Therefore, in many cases, there is a common cause, immediate or final, for both variables. Forest degradation can lead to a loss in the stature (e.g., height) or foliage/branch (crown) cover and vigour of individual trees or forest stands (Herold et al., 2011). Likewise, some tree species may be lost or replaced by those that are less favoured such as invasive species. Forest disturbance can either be human induced as a result of fuel wood gathering or selective logging events, amongst others, or natural events such as lightning strikes, wildfires and extreme climate (Schmidt et al., 2015). Degradation occurs when such events are repeated with limited opportunity for recovery (e.g., through succession, regrowth) but also as a consequence of adverse environmental conditions (e.g., drought, flooding, drainage of peat soils, outbreaks of pests or diseases including insects and pathogens, over frequent prescribed or management burning). Degradation can also occur where the capacity to regrow (e.g., through loss of seed banks, fragmentation) is reduced (Wilson and Sader, 2002; Herold et al., 2011).

Drivers of Deforestation and Forest Degradation

There are a number of direct causes of deforestation and forest degradation in Tanzania, each dominant in different parts of the country, depending on common practices and the kind of forest dominant there. Miombo woodlands are mainly disturbed by shifting agriculture and agricultural expansion, charcoal and firewood demand for domestic and industrial use, logging, illegal and unsustainable harvesting of forest products such as logging of high-value timber, forest fires, infrastructural development, village expansions and resettlements (Sunseri, 2009). Other studies have highlighted the effect of the introduction of alien and invasive species to the savanna and miombo woodlands. These causes are indirectly driven by market and policy failures, rapid rural settlement expansion and urbanization (population growth) and poverty (United Republic of Tanzania, 2013a).

Drivers of deforestation in sub-Saharan Africa are embedded in socio-economic relations between the state, private sector, and local people, as well as in chronic poverty. Between 1990-2000 Tanzania's deforestation rate was over 400,000 ha or about 1.1% per annum (Chiesa et al., 2009). Deforestation takes place mostly on open-access public land while degradation occurs in village lands and in state forests. In 2004, timber exports were banned as a result of a scandal involving illegal logging and corruption (Chiesa et al., 2009). Then in 2005, a report revealed the poor state of governance in the forestry sector in the country (Mustalahti et al., 2012).

Tanzanian forest area is declining as a result of agricultural expansion, uncontrolled wild fires, intense livestock grazing, illegal mining and charcoal making (**?**Burgess et al., 2010). A fieldwork to Liwale, in Lindi region, southern Tanzania concurred with most literature, that forest degradation was the main contributor to GHGs emissions than deforestation in the country. It was observed that among the leading drivers of forest degradation is shifting cultivation. Subsistence agriculture, which is predominantly shifting agriculture, is the most important direct driver of forest disturbance in Tanzania. Fuel wood collection, charcoal production, and, to a lesser extent, livestock grazing in forests are also major direct drivers of degradation, not just in Tanzania, but also in large parts of Africa (Kissinger et al., 2012). Wood harvesting for timber extraction is also a common driver observed in Liwale.

The main indirect driving forces of forest change are all expected to increase in the coming years. These include population and economic growth based on the export of primary commodities, national and international demand for agricultural products (food and biofuels), wood products and minerals are all expected to increase in the coming years (Sunseri, 2009; United Republic of Tanzania, 2013a).

Spatial Extents and Patterns of Degradation

Degradation may be localised (e.g., involving the loss of individual trees or parts of trees) or widespread (e.g., through wildfires covering thousands of hectares). Patterns of degradation vary from selective removal of individual trees or groups of trees, with the latter often leading to the division of forests into fragments that are often more susceptible to degradation (e.g., as a consequence of edge effects). In Tanzania, selective logging has been recorded as one of the major sources of degradation, and accounting for this kind and scale of disturbance is very hard. Another relatively common driver in south-eastern Tanzanian miombo forests is shifting cultivation. This

is a rife practice, as detailed under the Fieldwork Section in Chapter 4. This practice was found to be key in shaping the landscape in Liwale.

Temporal Dynamics of Degradation

The degradation of forests can take place after a single disturbance event or through a gradual and incremental process of depletion. Where a temporary loss of forests occurs, whether over a year or decades, these are not regarded as degraded. Nevertheless, forests may be degraded even though regrowth is occurring (e.g., in gaps left behind following removal of trees). Temporal thresholds for each forest and forest degradation type should be defined to avoid combined effects of short and long-term reductions in carbon stocks (e.g., selective logging followed by intensive logging) (Sasaki and Putz, 2009; Matricardi et al., 2010; Kissinger et al., 2012).

Measures of Degradation

When assessing degradation, measures relating to vegetation stress and damage to the forest canopy are often used. Changes in forest types can occur (e.g., in terms of dominant species composition) where new environmental conditions lead to a change in their floristic composition (Hansen and Loveland, 2012).

2.5.2 Detecting Degraded Areas

Remote sensing estimates of the area of degraded forest or the processes of degradation can be generated using either a direct or indirect approach. Herold and Román-Cuesta (2011) observed that the process of measuring and accounting for forest degradation and related carbon stock changes (losses) was more complicated and costlier than deforestation. The normal practice among nations faced with measuring degradation has been to use field data measurements, or use remote sensing approaches, or use a combination of the two which according to Herold et al. (2011), provide the strongest estimates. Unfortunately, most developing countries lack historic field data. The few that have such data do not have it consistently, which limits countries ability to integrate both approaches for the task. It results in most of the developing nations relying strongly on the remote sensing approaches. In other countries, even the available field data are not easily accessible, which further complicates things for individuals and organisations undertaking research

there. In Tanzania, for example, one of the conditions for access to the NARFOMA field data is that the spatial location of the plots can not be displayed, which introduces a limitation to a remote sensor who would have integrated it, or used it for training the classifier and for validation purposes.

This highlights the need for the advancement of remote sensing methodologies for developing countries. Work that has been done that advance the use of remote sensing to detect forest disturbance, and the commonly used remote sensing-based approaches for measuring degradation are either direct and indirect.

Direct approach

The direct approach compares satellite-derived measures of biophysical attributes (e.g., canopy cover) over time to indicate the process of degradation or identify events. Direct detection of degradation processes and related area changes focuses on forest canopy damage. The features enhanced and extracted from the satellite imagery are forest canopy gaps, small clearings and the structural forest changes resulting from disturbance (Herold and Román-Cuesta, 2011). A limitation of this approach is that often there is insufficient knowledge of properties such as the structure of undisturbed forests. Examples of studies that employed this approach either on a single date or composite image include Souza et al. (2005), Asner et al. (2005) and Oliveira et al. (2007). Other researchers use the direct approach of detecting the direct degradation indicators through a time series data analysis, and examples include Lambin (1999) and Healey et al. (2005).

An alternative method under the direct approaches is to extract combinations of optical and radar data from known areas of undisturbed forests within discrete ecoregions and then compare measures (e.g., backscatter, reflectance) from objects of unknown class (Herold and Román-Cuesta, 2011). Through this approach, degradation is based on differences in remote sensing data relative to those extracted for the undisturbed state. Such an approach requires knowledge of areas of undisturbed forests which can be achieved through reference to time-series of satellite-sensor data or aerial photography. For some areas/regions though, there may be no undisturbed forests, and in that case, the comparison is made with forests that are considered to be as close as possible to the undisturbed state (Kissinger et al., 2012). Examples of studies that used this approach include Jubanski et al. (2013), Goh et al. (2013), Dusseux et al. (2014) and Miettinen et al. (2014).

Indirect approach

The indirect approach extract features (secondary indicators) relating to infrastructure (e.g., roads, settlements, log landings) that might be used as proxys for newly degraded areas (Dons et al., 2016), and assuming, or more precisely, modelling degradation buffers around them (Potapov et al., 2008). One considers distance measures to disturbance events or infrastructure on the assumption that forest in closest proximity to these will be at a higher risk of deforestation and/or of continued degradation because of greater accessibility to humans (FAO, 2006b). Thereafter, one stratifies forests by ecosystem type into low, moderate and high risk areas based on statistical models that classify the risk of deforestation based on distance from already deforested areas and other geographical determinants (Herold et al., 2011). Even though this approach is noted to be useful for detecting newly degraded forest areas, it is less effective for detecting activities that are repeated and leading to degradation (Herold and Román-Cuesta, 2011).

This approach can be broken down to either the 'intact forest' or the 'deforestation modelling. In the intact forest approach, the presence of human infrastructure is viewed as a proxy for degradation, and therefore, its absence is used to identify intact forest land, without anthropogenic disturbance intact forest (Herold and Román-Cuesta, 2011). Studies that have employed this kind of an indirect approach include Mollicone et al. (2007) and Potapov et al. (2008). With the deforestation modeling approach, it can be applied to estimate both future and historical forest degradation dynamics through scenario modelling for forest degradation. With the right support from field data, similar modelling can be further used to (re)constructing historical and future scenarios of forest degradation (Herold and Román-Cuesta, 2011). Examples of studies that employed this approach include Grainger (1999), Soares-Filho et al. (2006) and Dons et al. (2016).

2.5.3 Calculating Emissions from Degradation

Emission factors resulting from forest disturbance are based on the area of forest affected by degradation (activity data), the change in forest carbon stocks between two dates. In an ideal environment, consideration is given to transitions from low to high risk (i.e., degradation) as opposed to high risk to another land use (deforestation). Remote sensing has been used extensively for environmental studies and for monitoring purposes. Whilst optical sensors such as Landsat have been around since 1970s and have been used for deforestation monitoring in the past, the

recent provision of space-borne data at higher spatial and temporal resolution (e.g., WorldView, RapidEye) has provided new opportunities for forest monitoring, particularly in relation to the detection of degradation and disturbance. Wang and Zhou (2011) lists some of the new opportunities and options as being cheaper, macro, fast, and enabling ease of dynamic and long-term monitoring. However, application can be limited owing to national restrictions, time and high costs associated with some satellite data (Lucas et al., 2007).

2.5.4 Status of Tanzanian Forests

Tanzania is still one of the least urbanized countries in Africa (United Republic of Tanzania, 2002). The main features of population distribution are: sharp discontinuities in density, with a number of densely populated areas separated from each other by zones of sparse population; the comparatively low population density in much of the interior of the country; and in most parts of the country, rural settlements tend to consist of scattered individual homesteads rather than nuclear villages. The population involved in agriculture has traditionally settled in areas suitable for crop production and mixed farming. Indigenous knowledge of trees and grasses was used as an indicator of land suitability. Rainfall and soil fertility are still decisive factors governing population distribution and density. About 10% of the country receives adequate rain (over 1000 mm per annum) and carries 60% of the population; 8% is fairly well watered and carries 18% of the population; 20% is poorly watered and carries 18% of the population; and 62% is very poorly watered and carries 1% of the population. Thus, about four fifths of Tanzania's population today is concentrated on only one fifth of its land. The rapid population growth is an environmental concern because of several reasons which include, threatening what is already a precarious balance between natural resources and people; shortening of fallow cycles, exhausting soil nutrients in agricultural activities; and increasing the demand for food and services, and consequently land (United Republic of Tanzania, 2002).

The 94.5 million hectares of land area in Tanzania is of a tropical climate, with 10 ecological zones that have different physiographic zones. Of the total land area, about 38% is covered by forests and woodlands. The country is categorised as one of the 34 biodiversity hotspots in the world. About a quarter of the land area of Tanzania is occupied by crucial ecosystems (e.g. national parks, game reserves and forest reserves) (Division of Environment, 2012).

There are two broad forest categories in Tanzania, these are reserved and non-reserved forests.

Reserved forest account for about 37% (12.5 million hectares) while unreserved forests account for 57% (19 million hectares) (United Republic of Tanzania, 2002). Reserved forests cover 37% land area, comprising of central and local government forest reserves, government-owned industrial plantations, and village land forest reserves (VLFRs) at the community level that was gazetted by the central government. Unreserved forests are village lands where forests and woodlands are not formally classified as reserves, and these cover 57% of the forest area (Mukama, 2010). Unreserved (unprotected) community lands experience heavy pressure from agricultural expansion, livestock grazing, wild fires, overexploitation, and unsustainable utilization of wood resources and other human activities (United Republic of Tanzania, 2002). Woodlands consist of just more than 96% of Tanzanias total forests, a majority being miombo.

The major use of wood in Tanzania is for fuel; hence about 95% of the country's energy supply is met from wood. For this reason, there are high rates of deforestation and degradation in both reserved and unreserved forests. Between 2005 and 2010, high rates of deforestation led to a loss of 403,000 ha of forest per year which was equivalent to 1.16% of forest area (United Republic of Tanzania, 2013a). Apart from deforestation and degradation, there is growing evidence that climate change is impacting on forests and forest ecosystems of Tanzania, and therefore livelihoods of forest dependent communities as well as national economic activities that depend on forest products and services are heavily stressed. Under a warmer climate forest ecosystems may also shift their ranges and lose some of their biodiversity.

Tanzania is a globally recognized storehouse of forest biodiversity, with two distinct forest-based global hotspots of high biodiversity found there. These are the eastern afromontane and the coastal forests hotspots (Burgess and Lovett, 2000). Even though the miombos are not a hotspot, SJP (2009) notes that they, together with acacia woodlands are also parts of high biodiversity wilderness areas that are renown for supporting intact assemblages of megafauna on the planet. These animals which are the footprint of the African landscape and known to require intact ecosystems for their conservation include large herbivores (e.g. elephants, rhinos, hippos, giraffes, buffaloes), migratory plains game (e.g. zebras, wildebeests, elands, gazelles), large predators (e.g. felids, canids, hyaenids, crocodiles, pythons), and large/migratory avifauna (e.g. vultures, raptors, ostriches, bustards, cranes, storks) (SJP, 2009).

Literature highlight degradation, deforestation and desertification as among the major environmental challenges faced in Tanzania. Recent droughts affected marginal agriculture leading to

soil degradation and erosion. Major direct causes of deforestation and forest degradation in the country include charcoal and firewood demand for domestic and industrial use, illegal and unsustainable harvesting of forest products, forest fires, agricultural expansion, overgrazing and nomadic pastoral practices, infrastructure development, settlement and resettlement, and invasion by alien species. These causes are notably indirectly driven by market and policy failures, rapid rural settlement expansion and urbanization (population growth) and poverty (United Republic of Tanzania, 2013a).

Sectoral Policy and Legislation

Tanzanian forests are protected by the Forest Policy, which was approved by the government of Tanzania in 1998 (Scheba and Mustalahti, 2015). Its overall goal is to enhance the contribution of the forest sector to the sustainable development of the country and the conservation and management of natural resources. It is focused on forest land management, forest-based industries and products, ecosystem conservation and management, and institutions and human resources (Institute of Resource Assessment, 2009).

The Division of Environment (2012) list the objectives of the forest sector to be the following:

1. Ensure sustainable supply of forest products and services by maintaining sufficient forest area under effective management.
2. Increase employment and foreign exchange earnings through sustainable forest-based industrial development and trade;
3. Ensure ecosystem stability through conservation of forest biodiversity, water catchments and soil fertility; and
4. Enhance national capacity to manage and develop the forest sector in collaboration with other stakeholders.

2.5.5 Miombo Woodlands

According to Mukama (2010), miombo woodlands are the most extensive tropical seasonal woodland and dry forest formation in Africa. They cover about 2.4 million km ^{2} of southern Africa (Mustalahti et al., 2012), stretching across Angola, Zimbabwe, Zambia, Malawi, Mozambique, Tan-

zania, most of the southern part of the Democratic Republic of Congo, as well as the northern parts of Namibia, Botswana and South Africa (Frost, 1996; Mpingo Conservation and Development Initiative, 2013), as shown in Figure 2.3. These woodlands play an important role for both rural and urban populations. Miles et al. (2009) highlighted that they support 87% of rural livelihoods, 90% of the national energy supply, and 75% of construction materials in the country.

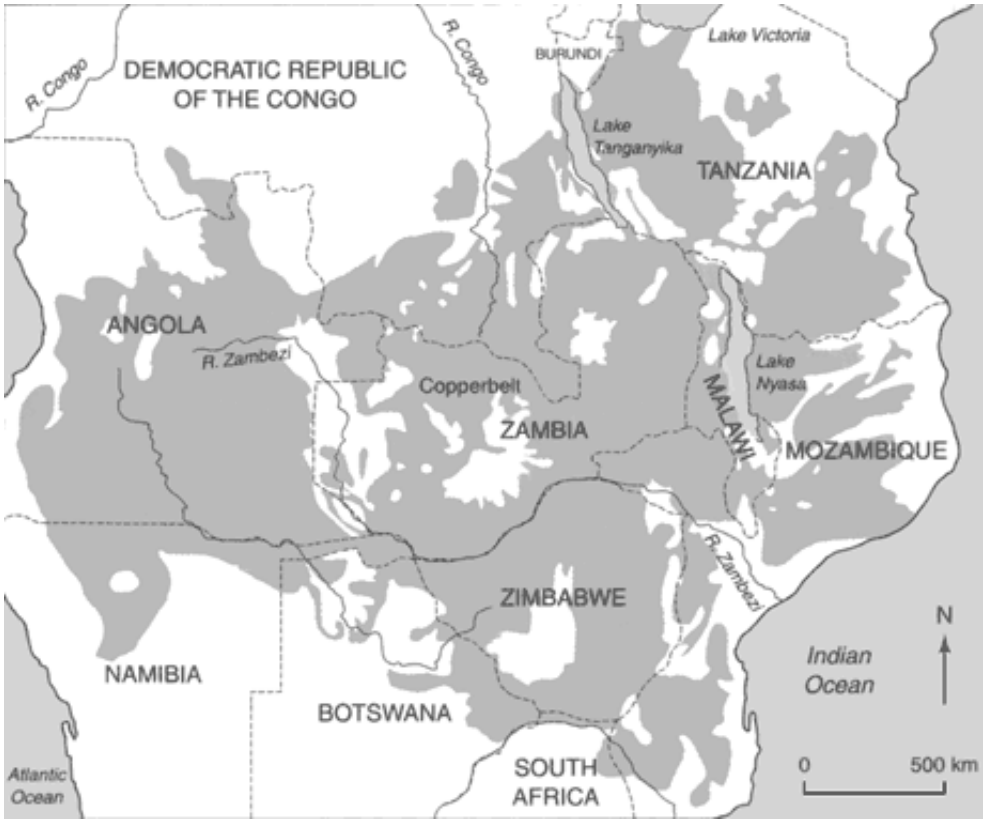


Figure 2.3: The spatial distribution of Miombo woodlands in southern Africa.
Source: Mpingo Conservation and Development Initiative (2013)

Lawton (1980) and Chidumayo (2013) state that the miombo is a vegetation type that has been maintained by man through a long history of cutting, cultivation and frequent dry season fires. These woodlands are characteristically dominated by a few trees, dominantly those in the subfamily *Caesalpinoideae*, especially species in the genera *Brachystegia*, *Julbernardia* and *Isoberlinia* (Campbell et al., 1996; Mukama, 2010). They usually have a shrub layer which varies greatly in its percentage cover, density and species composition, even though commonly dominated by *Combretum* and *Diplorhynchus* species (Mpingo Conservation and Development Initiative, 2013).

Miombos are commonly found in areas which receive more than 700 mm mean annual rainfall and nutrient-poor soils (Frost, 1996). They are semi-deciduous in nature; losing some or all of their leaves during the dry season, depending on the severity of the season. These woodlands provide resources that are vital to the livelihood of millions of rural and urban people living in and around them in central, eastern and southern Africa (Figure 2.3). In fact, about 40 million people obtain from these woodlands a multitude of products including food, energy, shelter, medicines and a number of invaluable environmental and spiritual services (Campbell et al., 1996; Abdallah and Monela, 2007).

Woodlands make up about 90% of all forested land in Tanzania and these are dominated by miombo, equivalent to 44.6 million ha, out of which 54% is under general lands (United Republic of Tanzania, 2001). The main concentration of miombo woodlands in Tanzania are found in Tabora, Rukwa and Kigoma regions (the western zone) and Iringa, Lindi, Mtwara and Ruvuma regions (the southern zone). The dry miombo woodlands occur in areas receiving less than 1000 mm rainfall annually including Shinyanga, Kigoma, Tabora, Rukwa, Mbeya and Iringa regions. On the other hand wet miombo woodlands occur in areas receiving more than 1000 m rainfall per year and these are found in south-eastern regions including Lindi, Mtwara, Songea, Mbeya, Iringa and Morogoro Campbell et al. (1996); Dondeyne et al. (2004). Miombo in Tanzania consists of two main layers, the tree canopy and the herb or ground layer, plus an under-wood layer of smaller trees. Some areas have a shrub layer, with the canopies of mature miombo stands reaching a height in the range 10-20 m. Some areas may be slightly open, otherwise most areas have a ground layer dominated by *Hyperrhenia* grasses and the saplings of the main canopy species. The canopy layer is often subjected to burning during the dry season (Mukama, 2010).

Fuller (1999) concluded that, from an ecological perspective, most miombo woodlands had been disturbed. There is very little 'old growth' woodland remaining. In many places the structure of the woodlands has been substantially modified through shifting agriculture, harvesting for firewood and charcoal, too frequent burning and, in drier areas, heavy grazing by livestock and wildlife (game). With repeated disturbances, closed-canopy miombo forests may degrade to open woodland and secondary grasslands. However, according to Chidumayo (1994), miombos may be restored through excluding fires and minimizing other disturbances. Being seasonal in nature, Figure 2.4 is the progression of a miombo forest through the year, showing it losing all its canopy cover during the dry season and achieving full canopy cover during the wet season.

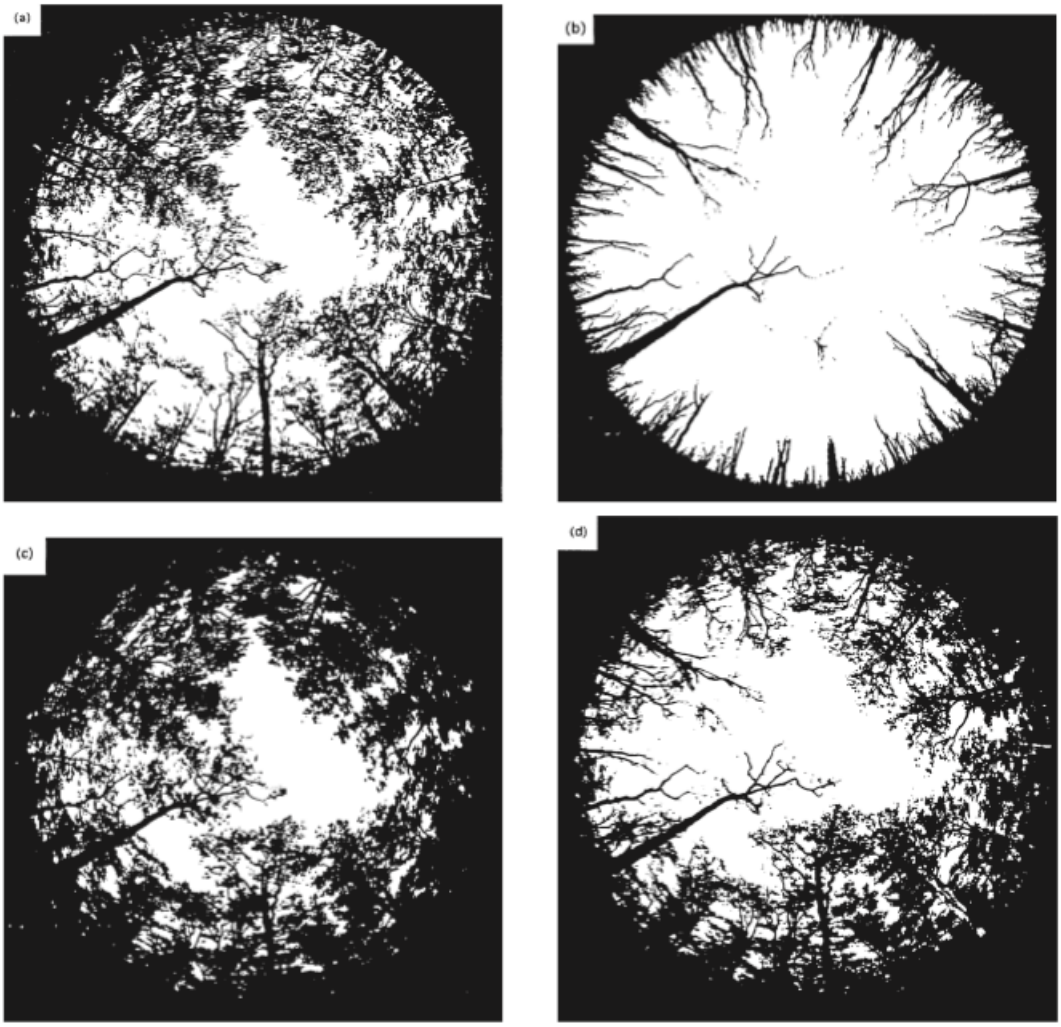


Figure 2.4: Black and white images showing the progression of tree canopy cover for a miombo woodland site in Zambia, taken using a flatbed scanner. The respective photographs were taken in (a) July 1988, (b) October 1988, (c) February 1989, and (d) May 1989.
Source: Fuller (1999)

A successful management of African miombo woodlands is crucial because these forests play a huge role in the sequestration of huge amounts of carbon, support livelihoods of millions of people in communities in these environments, provide a renewable source of energy (fuel wood and charcoal), and contribute towards poverty alleviation by supporting and strengthening local livelihood strategies (Dewees et al., 2011). Forests serve as a relief even in times of low food availability when they are a source of wild foods and fruits and other useful products (Paavola, 2008). Mustalahti et al. (2012) argued that if REDD+ was designed so as not to unduly restrict current forest uses for livelihood purposes, it had the potential to combine carbon sequestration with poverty reduction (Mustalahti et al., 2012).

Local Tanzanian studies that were undertaken to fill the gaps in the area of forest disturbance showed that there was substantial amount of disturbance in miombo woodlands, both from degradation and deforestation. For example, the Tanzania Forest Conservation Group (2007) reported agricultural encroachment, cultivation in the forest understorey, forest clearance for cultivation, timber harvesting, livestock grazing, pole cutting, firewood collection, hunting, gold mining, fire and charcoal production as the disturbances taking place in Nguru South, Mkindo and/or Kanga Forest Reserves in the country. It was concluded that amongst these, the level of disturbance caused by cultivation, hunting and timber harvesting had reached critical levels and that urgent action was needed.

2.5.6 Measures Aimed at Sustainable Tanzanian Resources Management

The overall national objective of Tanzania is to conserve natural ecosystems with their genetic resources so that the values and benefits of the forests are perpetuated, and not degraded. Through the UN-REDD+ programme, it aims to promote better management of natural resources and the environment through appropriate adaptation and mitigation strategies including REDD+ mechanisms (United Republic of Tanzania, 2012). The 1986-1988 Forest Policy which was fairly generalised such that it did not include specific policies that encouraged sustainable utilization of resources has been addressed since 2000 which saw the introduction of several policies such as the Forest Act No. 14 of 2002 (Government of Tanzania, 2002) and the Tanzania Forestry Action Plan (TFAP) 1991-2008 (Mitawa and Marandu, 1995) that support forest management. These policies have been enacted to specifically address issues like increasing awareness on nature conservation,

sustainable conservation by management of natural forest, research and training in conservation and creation of a network of nature reserves (Mitawa and Marandu, 1995; SJP, 2009)

2.6 Remote Sensing of Forest Mapping and Monitoring

2.6.1 Introduction

The assessment and monitoring of the forests of the world is a key for global change research, presenting valuable information for understanding the natural and man-made environments through quantifying vegetation cover (Lambin et al., 2001; Hansen et al., 2013; Dons et al., 2016; Joshi et al., 2016). This could be either at a local, regional or global scale (spatial), and either at a given time point or over a continuous period. Egbert et al. (2002), He et al. (2005), Wasige et al. (2013), Hansen et al. (2013), Alqurashi and Kumar (2013) and Joshi et al. (2016) are among studies and reviews that observed that knowledge on the current state of forests is critical as it informs decision-making and actions to be taken with regard to vegetation protection and restoration programs.

Traditional methods (e.g. field surveys, literature reviews, map interpretation and collateral and ancillary data analysis), however, are not effective to acquire vegetation covers because they are time consuming, date lagged and often very expensive. The technology of remote sensing offers a practical and economical means to study vegetation cover changes, especially over large areas (Næsset and Økland, 2002; Lippitt et al., 2008; Myint et al., 2011; Dupuy et al., 2012). The potential of remote sensing to systematically observe a landscape at various scales and times makes it suited for resources management. For some sensors such as Landsat, the technology extends possible data archives from present time to over several decades back. This has led to efforts by researchers and specialists to use earth observation data to map and monitor vegetation cover across different landscapes and scales (from local to global scale). Examples of global remote sensing-based products include the Global Land Cover Characterization (GLCC) Database, derived in 1992 from the Advanced Very High Resolution Radiometer (AVHRR) data (USGS, 2008), and the Global Land Cover 2000 (GLC2000) that was derived from SPOT4-VEGETATION imagery (GLC, 2015).

The latest land cover product at a global scale is by Hansen et al. (2013) that is a forest change

product, derived from Landsat data. There are other numerous products that have been derived even at the regional and local scale in other parts of the world, but not so commonly in the poorly researched African continent. For the Tanzanian savanna woodlands, the global products were found to be relatively generalised and with a coarse resolution that was not well suited to the local forest definition for these relatively open woodlands.

2.6.2 Remote sensing sensors

Simply put, a remote sensing sensor is a device that captures data about an object or scene remotely (without being in contact with it). Lillesand et al. (2008) defines it as the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation. Ground features/objects tend to have unique spectral features (i.e. reflectance or emission regions), and therefore can be identified from remote sensing imagery according to this unique spectral characteristics (Xie et al., 2008). Remote sensing imagery is acquired by a range of airborne and space-borne sensors, from multispectral sensors to hyperspectral sensors. This can be done with wavelengths ranging between visible and microwave, with spatial resolutions ranging from sub-metre to kilometres. The acquisition, depending on the sensor, is done with temporal frequencies ranging from a repeat cycle of only 30 minutes to a lengthy period of months (Xie et al., 2008; Joyce et al., 2009; Joshi et al., 2016).

Different sensors have different spatial, temporal, spectral and radiometric characteristics (Joyce et al., 2009). The selection of an appropriate sensor is important in vegetation cover mapping and monitoring as it determines how well the task can be performed. To correctly select images from an appropriate sensor, the mapping objective, the cost of images, the climate conditions (e.g. atmospheric conditions and season of the year where seasonality is an issue), and the technical issues for image interpretation are the four key factors that need to be considered (Xie et al., 2008; Cho et al., 2012; Cunningham, 2014). The mapping objective guides the remote sensor on what need to be mapped and to what accuracy. For example, high resolution imagery is best suited for fine-detailed classifications of vegetation, while low resolution imagery is usually best suited for the classification for broader land classes. High resolution satellite data are costly, a fact that need to be considered during the selection of the appropriate sensor for the task. The cost has to be considered against the spatial and temporal data needs for the area being studied.

The size of the area being mapped/monitored also influence the selection of the sensor to be used. Normally, vegetation mapping at a small scale requires high resolution images, while low resolution images tend to be suitable for large-scale mapping. Some parts of the world are known to suffer from high and persistent cloud cover. An appropriate sensor should be able to provide appropriate (cloud-free) images as per need (Soudani et al., 2006). Technical aspects usually considered include image quality, preprocessing and interpretation. Commonly applied sensors in vegetation mapping/monitoring include Landsat, SPOT, MODIS, NOAAAVHRR, IKONOS and QuickBird (Xie et al., 2008).

IPCC (2003) gives a more comprehensive list of important criteria for selecting remote sensing data and products for the task of forest mapping and monitoring as a sensor with adequate land-use classification scheme, appropriate spatial resolution, appropriate temporal resolution for estimating of land-use and carbon stock changes, availability of accuracy assessment, transparent methods applied in data acquisition and processing, and consistency and availability over time. As stated, the spatial resolution of the sensor selected is crucial in meeting the mapping/monitoring requirements. Mapping for small features and small change features (e.g. degradation) cannot be successfully achieved by using a low resolution image, but would require a high resolution imagery.

Zahabu (2008) noted that forest degradation is more difficult, if not impossible, to identify, let alone quantify, from remote sensing images. Degradation is not generally visible in medium resolution satellite images because much of it occurs below the canopy or in small patches which cannot be picked up at this resolution (DeFries et al., 2007). High resolution images such as IKONOS may be able to detect clearings down to ten square metres. For example, Souza et al. (2005) and Asner (2009) were able to identify larger disturbances such as logging in rainforests using this system. However, by 2008, Zahabu (2008) believed that apart from the cost associated with them, both in terms of acquisition and purchase, the manpower involved in analysing images at this level of spatial detail is prohibitively costly, if available at all.

Existing global land cover products (e.g. MODIS, NOAA AVHRR, GLC2000 and MERIS) derived from remote sensing were all based on time series of coarse resolution satellite data. Bontemps et al. (2012), Giri et al. (2013), Gong et al. (2013) and Ban et al. (2015) are among studies that highlighted the increased need for high spatio-temporal resolution data products. Ban et al. (2015) further stressed the emergence of the requirement for finer resolution land cover mapping at the global scale for studies in different fields such biodiversity, food security and forest carbon

studies.

As of 2008, the United States of America, through the National Aeronautics and Space Administration (NASA) and the Department of Interior United States Geological Survey (USGS) (Irons et al., 2012; USGS, 2015b), adopted a policy that made the past, present and future Landsat holdings available on a free and open policy basis (Woodcock et al., 2008; Belward and Skøien, 2015). More satellites have followed suit since then, with examples including China Brazil Earth Resources Satellite (CBERS) as of 2011 (ESA Earth Online, 2016) and, notably, the European Union (EU) in 2013 (European Union, 2013). The EU adopted the same policy for data at resolutions coarser than 10 m from its current and future Sentinel Earth observing missions (ESA, 2016; European Union, 2013).

A number of new Earth-observing missions have been launched for this decade, with more being planned. As a continuity to the already existing rich archive of the Landsat data time series, Landsat 8 (formerly the Landsat Data Continuity Mission, LDCM) was successfully launched in 2013, and is being taunted as the future of Landsat satellites, collecting valuable data and imagery for land cover mapping and monitoring studies (NASA, 2014; Roy et al., 2014). The launch of Sentinel-1 in 2014 and Sentinel-2 in 2015 (ESA, 2012, 2016) has also availed more remote sensing data that is suitable for forest mapping and monitoring. The high global acquisition frequency and distribution policy for products from these satellites makes them suitable even for near real time land cover mapping and monitoring, and will improve data availability, both the spatially and temporally (Simonetti et al., 2014).

Higher resolution data (less than 10 m) have continued to be distributed on a commercial basis. However, Belward and Skøien (2015) noted that some of these data are available for free viewing via web services, and an example is Google Earth. Even though a researcher may not have the privilege of accessing both the entire archives and original digital data, it provides land cover and cover change studies with a valuable source of information for qualitative assessments of cover change. Moreover, such data provides a crucial source of information for independent accuracy assessments and classification validation (Schneider et al., 2009; Giri et al., 2011; Belward and Skøien, 2015).

2.7 Conclusion

The chapter focused on the review of literature, touching on key topical issues for the study. These include an overview of the REDD+ initiatives; its role, carbon monitoring the Tanzanian REDD+ status quo, the finances it has availed, and ongoing REDD+ pilot projects, documenting lessons learnt from these projects. The chapter further conceptualised terms, and presented the working definitions for the study. It expanded on the subject of forests and their disturbance; the processes and causes of the disturbances, the drivers of the disturbances, the kinds of disturbance, and the techniques for detecting and measuring the disturbed areas. The chapter further presented the status of Tanzanian forests, and specifically the savanna woodlands that make up the study site. Lastly, the chapter documented the role played by remote sensing for forest mapping and mapping, presenting examples of studies that used remote sensing data.

Chapter 3

Description of the Study Area

3.1 Geographic and Physiographic Description of Tanzania

3.1.1 Location and Administration of Tanzania

Tanzania is located in the east African coast, falling within the sub-Saharan region and entirely within the tropics (Figure 3.1). It lies between longitudes 29^o and 41^o east of the Greenwich Meridian and latitudes 1^o and 12^o south of the Equator (FAO, 2006b). It has a total area of 945,087 km^2, of which 886,037 km^2 is land area and 59,050 km^2 is covered by water bodies. It is bordered by the Indian Ocean to the east; Kenya and Uganda to the north; Rwanda, Burundi, and the Democratic Republic of Congo to the west; and Zambia, Malawi and Mozambique to the south. On land, it is bordered by three of the largest lakes in the African continent; namely Lakes Victoria, Tanganyika, and Nyasa (Mitawa and Marandu, 1995).

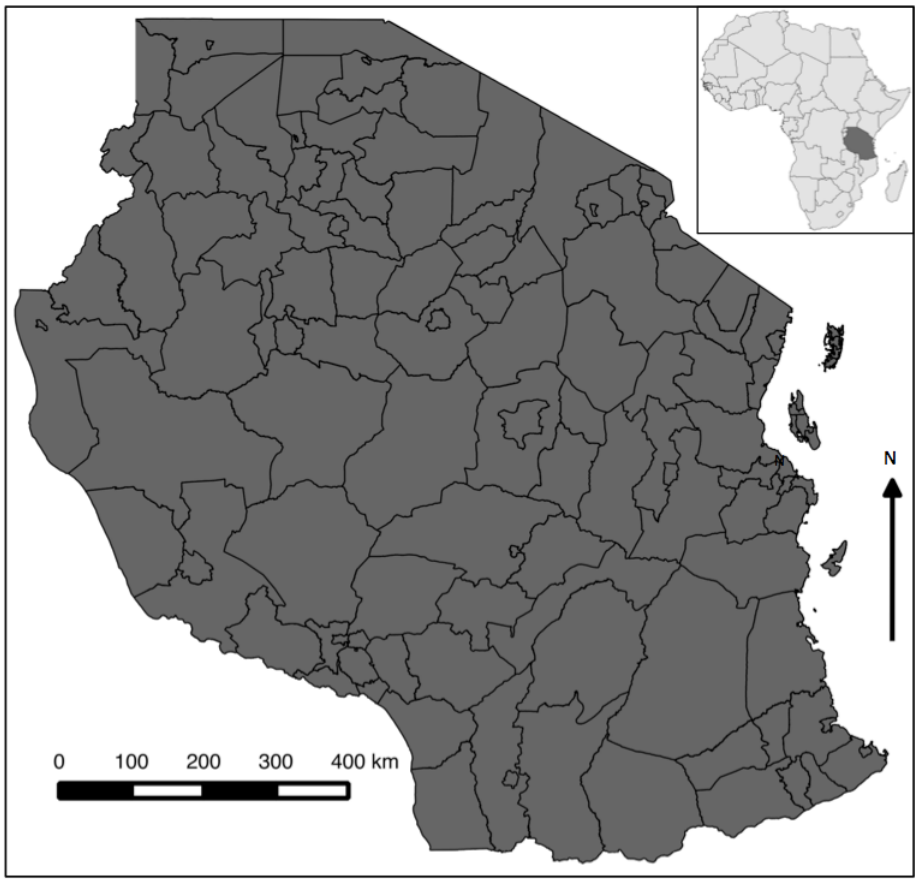


Figure 3.1: Map of Tanzania with the 130 districts making up the country. An insert shows its location in the African continent.

Administratively, Tanzania is made of Tanzania mainland (formerly Tanganyika) and Zanzibar, forming the United Republic of Tanzania (Mitawa and Marandu, 1995; FAO, 2006b). The mainland covers 881,289 km^2 (99.72%) of the land area, while Zanzibar archipelago which is made up of Unguja and Pemba account for 2,460 km^2 (0.28%). Tanzania mainland is made up of 21 administrative regions, with each region divided into districts, 130 in total. These districts are further subdivided into divisions, wards and villages respectively. Road and railway networks in Tanzania are generally poor, making accessibility to most areas difficult (The Government of Tanzania and FAO, 2009).

3.1.2 Physiographic Features of Tanzania

Tanzania is renown for its unique landscape as both the highest and lowest elevations in the African continent are found there. These are the summit of Mount Kilimanjaro (5,950 metres above sea level) and the floor of Lake Tanganyika (358 metres below sea level) respectively (The Government of Tanzania and FAO, 2009). Except for the coastal belt, most of the country is part of the Central African plateau, the eastern arc mountain ranges such as the Ngorongoro Crater (biodiversity hotspot), and the southern and northern highlands which cut across the mainland to form part of the Great Rift Valley. These are mostly characterised by gently sloping plains and plateaux broken by scattered hills and low-lying wetlands. Even though there are no large rivers traversing the country, three great African rivers (River Nile, Congo River and Zambezi River) have their sources in Tanzania (FAO, 2006a). The country is well drained with numerous river basins such as Rufiji, Pangani, Ruvu, Great Ruaha, Malagarasi, Kagera, Mara, Ruvuma, and Ugalla River.

Figure 3.2: Major towns and cities of the Republic of Tanzania.
Source: Tanzania Natural Resources Forum (2012)

3.2 Liwale District and Study Site

The research sought to develop methods that would be robust and scalable to the wider Tanzanian savanna and general African savanna landscape. The study focused on a site in the Liwale District in Lindi Region, south-eastern Tanzania (Figure 3.3). The site believed to possess environments that would enable the development of such methods was found in Liwale district, Lindi region in the south-eastern part of the country. It covers about 16,000 km^2 (1,600,000 ha; Figure 3.3), and covers approximately half of the Liwale District. The northern part of the site is a game reserve (Selous Game Reserve) whilst the southern part is owned by local populations living in villages.

3.2.1 Geographical location of Liwale

The district of Liwale lies between $36^o50^{'}$ and $38^o48^{'}$ east of the Greenwich Meridian, and between 8^o and $10^o50^{'}$ south of the Equator. It is the largest of the six districts in Lindi region, bordered by the district of Rufiji in the north-east, Ruangwa and Kilwa in the east, Nachingwea and Tunduru in the south and Ulanga in the north-west (United Republic of Tanzania, 1997). Figure 3.3 is the district of Liwale in the Lindi region. The study site covers about half of the district of Liwale.

Liwale district is divided into 3 Divisions, 16 Wards and 42 Villages, and covers 3,838,000 ha (Mukama, 2010). Approximately two-thirds of the entire district is covered by the Selous Game Reserve (Figure 3.4). The reserve covers an area of 2,558,600 ha (67%), and is comprised of the western, northern and the north-eastern parts of the region. The remaining 1,279,400 ha (33%) is village land area, either settled on or set apart as a forest reserve. According to Sundström (2010), Ngindo people are the predominant ethnic group in Liwale, while other local groups include the Mwera, Yao, Ndonde, Makonde and Ngoni people.

Liwale District is a hub of biodiversity, with Dondeyne et al. (2004) recording 133 different tree species within the landscape. This landscape is predominantly characterised by sandy soils, which are deep and poor in nutrient contents (Mukama, 2010). The forest area contains large patches of dry miombo, closed dense forests, riverine and wet miombo forests with some high-value timber species (Campbell et al., 1996; Abdallah and Monela, 2007; Mukama et al., 2012). The Mpingo Conservation and Development Initiative (2013) noted that the district still had good forest cover.

Figure 3.3: Liwale District, within Lindi Region, as well as an insert of Tanzania, showing the spatial location of the study site in the country.

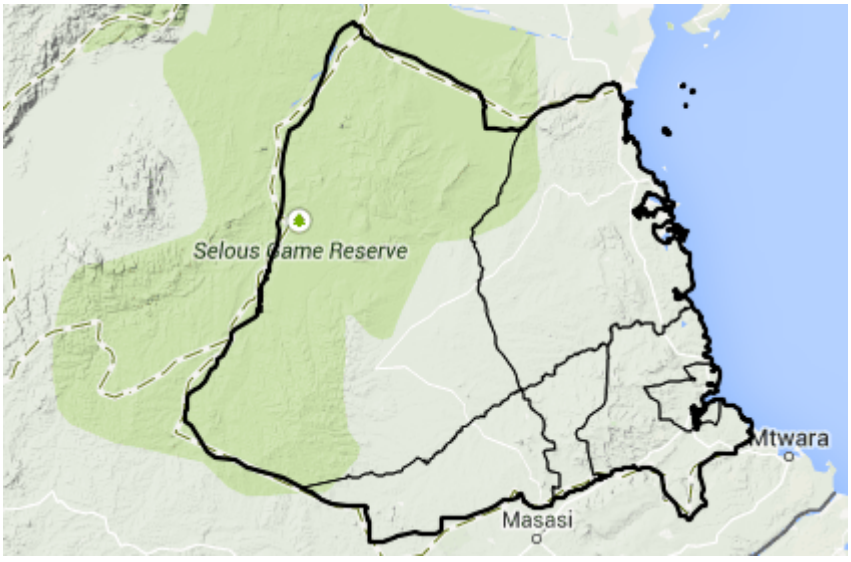


Figure 3.4: Liwale District and the spatial coverage of Selous Game Reserve.
Source: National Bureau of Statistics (2014)

However, logging pressures had hiked drastically in recent years.

3.2.2 Population

According to the 2012 National Population Census, the population of Tanzania (Mainland and Zanzibar) stands at 44,928,923 people, with an average growth rate of 2.9% (The United Republic of Tanzania, 2013). The average national population density currently stands at 44 persons per km^2. The Tanzanian population density averages about 44 persons per km^2. However, regions of Dar es Salaam, Kilimanjaro and Mwanza support over 200 people per km^2. On the contrary, some districts and regions such as Lindi, Rukwa, Ruvuma and Tabora are sparsely populated, with densities as low as 13 people per km^2 (The Government of Tanzania and FAO, 2009). Major reasons for the sparse population in these regions are poor infrastructural development, poor and unreliable rainfall and infertile soils which makes them not well suited for agricultural practices.

Table 3.1: The population of Tanzania, Lindi Region and Liwale District, and the total land area of each.

Coverage	**Population**			**Land area (ha)**
	Male	Female	Total	
Tanzania	21 869 990	23 058 933	44 928 923	88 580 000
Lindi region	414 507	450 145	864 652	6 604 000
Liwale district	44 027	47 353	91 380	3 431 400

Source: National Bureau of Statistics (2013)

According to the 2002 National Population Census, the Liwale District had a population of 75,548 people, a total of 38,742 women and 36,804 men (United Republic of Tanzania, 2003). Of the six districts that constitute the Lindi Region, Liwale was the least populated with a density of 2 people per km^2 (Table 3.1). The 2012 National Population Census found that the population of Liwale had risen to 91,380 (44,027 men and 47,353 women), a change of +1.96% per year (The United Republic of Tanzania, 2013). This population growth means that even though still the least populated in the region and amongst the least populated in the country, Liwale has seen a population density rise from 2 to 3 people per km^2 by 2012 (Table 3.2).

Table 3.2: Population statistics for Lindi Region and Liwale District, for the years 1988, 2002 and 2012.

Year	Lind Region			Liwale District		
	Population	Change (%)	Density (ppl/sq. km)	Population	Change (%)	Density (ppl/sq. km)
1988	646 494			52 240		
2002	787 624	+0.28		75 128		2.0
2012	864 652	+0.93	13.1	91 380	+1.96	2.7

Source: United Republic of Tanzania (2003) and National Bureau of Statistics (2013)

3.2.3 Climate

The country has a great diversity of climatic conditions which FAO (2006a) attributes to the proximity to the ocean and the numerous inland lakes, altitude (which controls temperature), as well as latitude. The mean annual rainfall varies widely from about 320 mm to over 2500 mm per annum (Mitawa and Marandu, 1995), depending on altitude and latitude. Even though it is a well watered country, there are considerable annual variations in the amounts and durations of rainfall. Most parts of the country experience long dry seasons, with the rainy season lasting from March to June. The central plateau is the most arid part, receiving the least rain (<600 mm per annum). The north-western highlands are cool and temperate with the rainy season lasting from November to December and February to May. Annual temperatures in Tanzania range from 24 ^{0}C to 34 ^{0}C, the minimal variations in monthly temperatures being attributed to its close proximity to the Equator (The Government of Tanzania and FAO, 2009).

The climate of the Liwale District is influenced by the southern-eastern winds in mid-year and the northern-eastern winds during the turn of the year. Local temperature ranges between 20 ^{o}C and 30 ^{o}C and the average temperature is 25 ^{o}C (Mukama, 2010). Liwale experiences a dry season between the months of July and October. The main rainy season is from March to May each year. However, before that season, there is a short rain period starting from late November to January, which makes the area have a bi-modal rainfall pattern. The annual rainfall for Liwale ranges from 600 mm to 900 mm (Frost, 1996; Tanzania Meteorological Agency, 2014).

3.2.4 Soil and Geology

The soils and geology of Tanzania are very diverse. Along the coast, soils range from deep, to sandy, to heavy textured soils with moderate to high available water content. Central and western plateau areas have got sandy loams of low nutrient content and low water holding capacity. Drought-

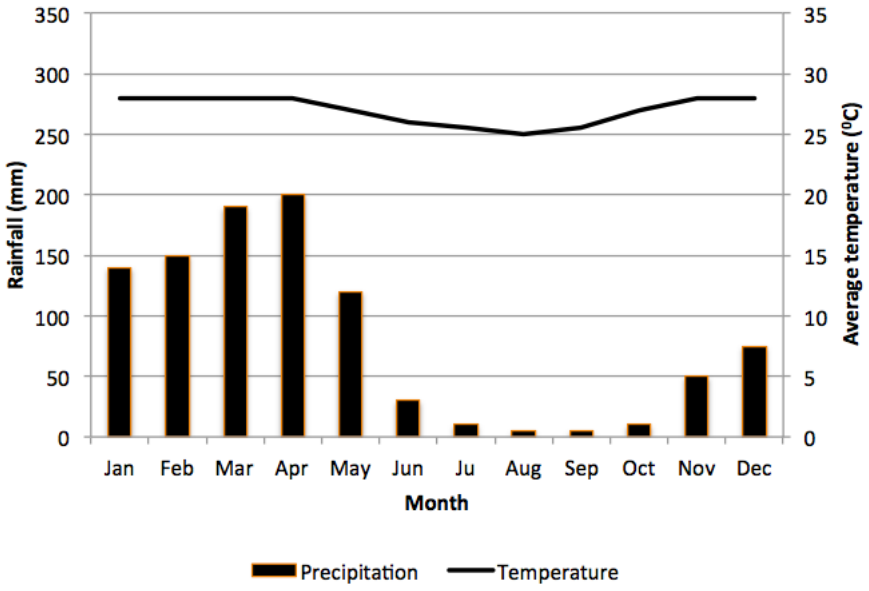


Figure 3.5: Annual temperature and precipitation for Lindi Region.
Source: Green Resources (2013) and Tanzania Meteorological Agency (2015)

prone soils are found in the north, in the Masai steppe and the south-eastern plateaus. The hill and mountain slopes and the central highlands have eroded and deeply weathered soils that are susceptible to erosion (United Republic of Tanzania, 2009).

The miombo woodlands have nutrient-poor soils (Frost, 1996), predominantly classified as hypoluvic arenesols. Profondic and arenic luvisols are also common within these environments (Mukama, 2010). Approximately 80% of the forest area is occupied by a sandy soil that is deep and poor in nutrient contents (Campbell et al., 1996). Dondeyne et al. (2004) reported that the soils in the undulating plains of Liwale have low fertility, with their top soil having a low organic carbon content (0.2 - 0.6%).

3.2.5 Forests and Woodlands

Tanzania has a total forest area of 35.3 million ha, 16 million ha of which are gazetted forest reserves. These reserves have been set aside by villages and are usually rich in biodiversity and contain trees that of high value. A further 2.2 million ha are protected by national parks and nature reserves. The remaining 17.3 million ha are unprotected and occur as unreserved forests or open access public land (United Republic of Tanzania, 2009). Mustalahti et al. (2012) noted that

miombo woodlands are the main forest type, covering 40% of the land area. These forests are unique natural ecosystems that provide a range of services including habitat for wildlife, genetic resources, and bees for apiculture. Left standing, the forests provide substantive stocks of carbon. They also contribute exports and tourism earnings to the country's Gross Domestic Product (GDP). Over 700,000 people are employed in various forest related activities.

The real contribution is under-estimated due to unrecorded labour in the collection of wood fuels and other forest related products consumed by households. This contribution forests make to human lives, whether directly or indirectly, is under-estimated because of unrecorded consumption of wood fuels, bee products, catchment and environmental values, and other forest products. However, it is estimated that the sector's contribution to the GDP is between 2.3% and 10% of the country's registered exports. The forest resource also offer services to other sectors such as pasture for livestock, raw materials for industries protection of watersheds, source of water for irrigation, generation of electricity, environmental protection, control of soil erosion and nutrients.

Two-thirds of Tanzanian woodlands are on public lands, and lack proper management (The Government of Tanzania and FAO, 2009). Moreover, since bio-energy is the main source of fuel for the rural population and accounts for 92% of the total energy consumption in the country (Miles et al., 2009), these are being rapidly exploited. Despite all the importance and roles played by the forest resources to the economy, there are a number of problems faced which hamper the development of the sector and and underestimate its contribution. The various problems include among others deforestation, inadequate forestry extension services, inefficiency wood based industries and poor infrastructural facilities. Others are outdated legislation, fragmented administration at all levels between the centre and the local levels, lack of participation of various stakeholders in the management of the resources and poor resource databases, and outdated and non-existence of management plans for efficient resource use (The Government of Tanzania and FAO, 2009).

Of the 1,279,400 ha village land outside the Selous Game Reserve, 13 villages gazetted a total of 139,420 ha of forest comprised of miombo woodland and high value tree species, as a forest reserve known locally as the Angai Village's Land Forest Reserve (AVLFR). Even though it is contiguous, it is still owned and separately managed by each individual village, and the ownership breakdown is given in Table 3.3. AVLFR is one of the largest areas in Tanzania to be under Participatory Forest Management (PFM) and since 2005, it has served as one of the national demonstration sites for carbon monitoring in preparation for participation in the REDD+ mechanism. The 13

villages received land certifications and titles for these in 2005. However, following a five month investigation in 2014, a report by the Guardian on Sunday newspaper revealed that rampant corruption had turned the woodland ecosystem into a 'Mecca for timber smugglers', reducing it to a 'mere shrub' through degradation (Kitabu, 2014). Of relevance is that national reports by The Government of Tanzania and FAO (2009) did not include degradation as one of the problems undermining the role of woodlands on the economy and livelihoods of local populations.

Table 3.3: Villages with AVLFR ownership in Liwale District.

Village name	**Forest area (ha)**	**Village population**
Nahoro	40 939	1 623
Nangano	2 601	712
Kibutuka	5 402	1 800
Kiangara	2 296	1 825
Kitogoro	7 425	1 103
Mtawatawa	11 761	1 077
Mikunya	1 628	1 683
Liwale B	7 135	5 759
Likombora	19 855	1 463
Mihumo	11 792	3 015
Ngongowele	8 285	2 320
Ngunja	6 626	1 186
Lilombe	13 675	3 390
Total	**139 420**	**26 956**

Source: The Government of Tanzania (2009)

Phenology and Cycle of Tanzanian Miombo Woodlands

Miombo is the Swahili word for *Brachystegia*, a genus of trees comprising a large number of species. Miombo woodland is classified in the tropical and subtropical grasslands, savannas, and shrublands biome (in the World Wildlife Fund scheme). The biome includes four woodland savanna ecoregions (listed below) characterized by the predominant presence of Miombo species, with a range of climates from humid to semi-arid, and tropical to subtropical or even temperate.

The African miombo woodlands are characterized by dry miombo, closed dense forest, riverine and wet miombos. Overall, they are rich in biodiversity as documented by Dondeyne et al. (2004) who identified 133 different tree species in 2004. However, the diversity of canopy tree species is low, overwhelmingly dominated by high valuable timber species in the genera *Brachystegia sp.*, *Julbernadia sp.* and *Pterocarpus angolensis* (Campbell et al., 1996; Ryan et al., 2010). Ribeiro et al. (2012) found that in mature miombos, these species comprised of an upper canopy layer made

of 10-20 m high trees. In some parts, there is a scattered layer of sub-canopy trees. There is also a discontinuous understory layer and comprised of a mixture of broadleaved shrubs, suppressed saplings of canopy trees, a sparse but continuous herbaceous layer of grasses, forbs and sedges, and bracken (*Pteridium aquilinum*) (Campbell et al., 1996). These forests are also important as they occur on or in proximity to the source of many rivers and streams that support both local populations and wildlife.

Forest monitoring requires the ability to separate forest from non-forest areas. The ability to do so is especially crucial where monitoring is based on remote sensing. In these woodlands, most of the trees are deciduous/semi-deciduous; losing some or all the leaves at the end of the wet season, and into the dry season, which is around the month of June. All the *Brachystegia* species are deciduous, with the exception of *B. taxifolia* which is evergreen. As already mentioned, Liwale experiences a bi-modal rain pattern, with the first short rain period starting from late November to January, and the main rains received between March and May. Phenology is the duration and intensity of leaf display, and provides a fundamental descriptor of vegetation types and function (Fuller, 1999). For example, Ribeiro et al. (2012) noted that changes in phenology helps us better understand the cycle of forests; the type of environment in which the plants survive, as well as their adaptive strategies to environmental factors (e.g. nutrient availability, water and solar radiation).

Peak leaf cover of trees but also the understorey (shrubs, grasslands and herbs) occurs during and especially towards the end of the wet season. Following leaf fall, the landscape is comparatively bare for 2-3 months (Shorrocks, 2007). However, some may remain on the branches for much of the dry season unless they are removed by strong winds or frosts, or destroyed by fire. These flammable systems are strongly influenced by anthropogenic fires (Ryan et al., 2012). Whether these are natural or human-induced, can be frequent due to the strong rainfall seasonality. Then a few weeks before the rains return, there is a flush of new leaves in the dry heat, often a shade of pink, reddish or brown (Shorrocks, 2007; Spinage, 2012), with many also flowering and setting fruit. Within about three weeks the leaves will turn green. Trees growing on shallow soils and in dry marginal miombo habitats may not come into new leaf until the first rains have occurred. Generally, over most of the miombo woodland, trees are in leaf at least two months before the start of the rains (Lawton, 1980).

Fire is an important part of the miombo landscape and can be natural or human driven, even though Sheuyange et al. (2005) noted an increase of those anthropogenic to over 70% in the

continent. Fires are most prevalent during the dry season and particularly at the end, when vegetation is driest. The savanna forests are notably adapted to the disturbance of fire, and (Käll, 2006) observed that frequent anthropogenic fires positively influence the tree species richness and create a more heterogenic savanna since repeated burning creates a mosaic of vegetation patches in different stages of recovery.

Literature on phenological properties and patterns of the understory vegetation (grasses and shrubs) is lacking. Whilst dieback during the dry season is indicated, the timing in relation to the overstorey has not been documented. Lack of literature in this regard is a limitation as a better understanding of both overstory (forest) and understory (non-forest) is critical in informing the best time for acquiring images when the two are separable based on phenology. A documentation of the actual times of the year this is the case is needed.

3.2.6 Drainage

The district does not have large rivers, most are seasonal and tend to dry up during prolonged dry seasons and drought periods. The major rivers include Liwale, Mbwemkuru, Angai, Mihumo, Nangula, Nakawale and Nchonda (W Charitable Foundation, 2006; Mukama, 2010). Angai Mbwemkuru and Mihumo rivers are the major sources of water for people in Nangano area. In some villages such as Ngunja, Ngongowele and Mihumo, there is a combination of deep and shallow water wells. However, the shallow wells tend to dry up during the dry season. During the dry season spells, some villagers travel up to 5 km to access water (Mukama, 2010).

3.2.7 Crop Production and Livestock Keeping

It is estimated that approximately 90% of the population is engaged directly or indirectly, in agricultural activities, which provide about 50% of GDP and more than 75% of foreign exchange earnings. It is estimated that more than 92% of the population adjacent to forest areas depends on agriculture and forest products for their livelihood (Mustalahti et al., 2012). Agriculture is the main economic activity,and is dominated by small-holder farmers practising rain-fed subsistence agriculture on farms with average sizes between 0.9 ha and 3.0 ha. The main food crops produced include maize, cassava, sorghum, rice and various legumes. Indigenous cultivation practices are still popular as about 70% of Tanzania's crop area is cultivated by hand hoe, 20% by ox plough

and only 10% by tractor (The Government of Tanzania and FAO, 2009). Major constraints facing the sector is the falling labour and declining land productivity as a consequence of poor technology and dependence on unreliable weather conditions. Women are still the core producers of food as they still constitute the main agricultural labour force. Farmers tend to extensively use the land closer to villages, and shifting cultivation is commonly practised to produce food crops, a practice that has a bearing on the local forests. Farmers also tend to have permanent fields of cashew trees and sesame, where intercropping is also practised to supplement food crops.

Livestock production contributes to national food supply as it converts rangelands resources into products suitable for human consumption and is a source of cash income for farmers. The most common form of livestock rearing is pastoralism which is practised at a subsistence level in traditional grazing areas. This is dominantly in the drought-prone northern plains where climatic and soil conditions do not favour crop production. The main roles of livestock in this system is to store wealth and also serve as a source of cash income. Commercial ranching exists in different regions of the country, and stock numbers per farm are generally low, accounting for only 2% of the total livestock in Tanzania (The Government of Tanzania and FAO, 2009).

Even though the miombo woodlands of Liwale are on nutrient-poor soils, there is barely any practice of pastoralism. According to Malele et al. (2011), southern Tanzania is one area in the country where livestock-keeping has been severely constrained by tsetse transmitted *trypanosomiasis* as the area is infested with tsetse flies. Tsetse flies are the vectors of *trypanosomes*, the causative organisms of *trypanosomiasis* and *nagana* in animals and sleeping sickness in man. Malele (2011) reported that about 40% of land suitable for grazing and areas with high agricultural potential are currently tsetse infested, and remain unexploited. Nationally, it is estimated that about 4.4 million livestock and 4 million people are at risk of contracting tsetse borne trypanosomiasis. This makes cattle rearing too difficult in Liwale District, and leads to loss of agricultural productivity. This ultimately forces the dominantly subsistence population to heavily rely on the natural resources (forests) as their main sources of livelihood.

3.2.8 Energy Resources

Biomass energy resource comprising of fuelwood and charcoal from both natural forests and plantations account for 93% of total energy consumption in Tanzania, while imported petroleum, hydropower and coal are the major sources of commercial energy in the country (Miles et al.,

2009). Only 0.6 % of the energy is from electricity generation and is driven by hydropower which is also dependent upon a consistent water supply. Hence, drought may impact on this generation for such a drought-prone environment. In an attempt to reduce the population's over-reliance on biomass energy, the government has worked towards sourcing electricity from neighbouring states, to boost the electric supply on the national grid. But because only urban areas are connected to the national grid, the move has not improved the situation on the over-exploitation of forest resources for energy by the rural communities. In 2009, the government acknowledged the urgent need to supply these rural areas made up of about 8,200 villages with electricity to curb deforestation (The Government of Tanzania and FAO, 2009). However, such programs will take time to come to fruition, and time being, forests continue to be depleted at alarming rates. Liwale district is dominantly rural, with most of the population still reliant on forest resources for energy in the form of fuelwood and charcoal. There are a number of villages within the district that still do not have access to electricity.

3.2.9 Tourism and socio-economic activities

In general, the road and railway network in Tanzania is poor and many areas are inaccessible. Tanzania is, however, well endowed with tourist attractions, including national parks and reserves, and the wildlife they possess. As such, this sector currently contributes 14 % of the GDP, and has an even greater potential. It currently employs over 30,000 people. The country's natural resources possess a great potential to support a strong economy; with extensive fertile agricultural land and pasture land, popular wildlife reserves, woodlands and forest ecosystems that are a global biodiversity hotspot. This again highlight the massive role the natural environment (forests) plays, and will continue to play in shaping other sectors that rely on it (The Government of Tanzania and FAO, 2009).

3.2.10 Wildlife

Tanzania has a rich and diverse range of fauna and flora including a wide variety of endemic species and sub-species. The biological diversity and degree of endemism consist of primates (20 species and 4 endemic), antelopes (34 species and 2 endemic), fish (with many endemic in Lake Victoria, Tanganyika, Nyasa and other small lakes and rivers), reptiles (290 species and 75 endemic), and

amphibians (40 endemic) invertebrates and plants (approximately 11,000 species including many endemic).

In an attempt to preserve this diverse floral and faunal species nationally, about 19% of Tanzania's land area is dedicated to wildlife in protected areas where no human settlement is allowed. Furthermore, in 9% of the country, wildlife and humans co-exist. Environmentally, this has gone a long way in preserving forest areas as there is minimal human-induced disturbance on these preserved areas. For example, as discussed above for Liwale, Selous Game Reserve ensures the continued preservation of plant and animal life in the region, which may not be as sustained in the open reserve areas around it. While the move enhances flora and fauna preservation, the country's rich wildlife is a unique natural heritage and resource, capable of boasting the economy through tourism. About 500 people are employed in this sector. The communities living adjacent to protected areas do benefit through hunting animals for game meat and get support of services from the private companies operating nearby and the government institutions related to wildlife sector (The Government of Tanzania and FAO, 2009).

The sector however, is undermined and constrained by illegal hunting (poaching) especially of the endangered species like elephants, competition with other land users, lack of public awareness of wildlife importance, lack of baseline data and information, inadequate rural user rights to the community and limited capacity in terms of budgetary allocation and human resources.

3.3 Conclusion

The chapter discussed the study site of Liwale, in Lindi region, southern Tanzania, in detail. Initially, the general geographic and physiographic description of the country were given. Then the properties of the district of Liwale, which has the site were discussed, with a focus on its geographical location in the country, its population, climate, soil and geology, forests and woodlands, its drainage, agricultural activities, as well as wildlife and tourism.

Chapter 4

Datasets

4.1 Introduction

This study used a number of remote sensing datasets, and also benefited from ground data collected over Liwale site during fieldwork in February 2014. The primary datasets for the study were RapidEye and Landsat satellite imagery. The other datasets the study used was LiDAR and Google Earth images, which were used for comparison and validation purposes. This chapter gives a summary of the properties of each dataset, then details how it was acquired, prepared and used in the research.

4.2 Remote Sensing Data

The earth has been and continues to be observed by many sensors mounted in different airborne and space-borne vehicles. Each sensor has been designed to meet certain requirements in line with the primary motivation behind its development. Requirements range from agriculture, forestry, land cover and land use, hydrology, oceans and coastal monitoring, atmospheric monitoring, to geological mapping studies (Lillesand et al., 2008; Cunningham, 2014). Airborne sensors usually acquire data a few kilometres above the surface, the most common being those mounted on aircrafts. Space-borne sensors greatly vary, some just outside the edge of the atmosphere while others are thousands of kilometres away. They have played a huge role in aiding communication, navi-

gation,defence, and lately scientific research. Even though the number is still relatively small, the last decade has seen a rise in sensors dedicated to the critical role of monitoring the atmosphere and the Earth surface for environmental conditions monitoring. A number of new Earth-observing missions such as Landsat 8, Sentinel-1 and Sentinel-2 (USGS, 2015b; ESA, 2012) are planned for this decade.

During the planning stage and the selection of data to be used, it is essential to first review sensors that could provide remotely sensed data for the study area. The review has to take into consideration several factors to ensure an appropriate sensor is chosen for a task. Factors considered include sensor coverage, pixel resolution, temporal resolution, spatial resolution, spectral range of bands, and the cost associated with data acquisition and products (Cunningham, 2014). This study aimed to explore the extent to which remote sensing data can map forests and detect change within them, meaning that the sensors considered had to be ones that are appropriate for forest monitoring. Even though not a lot of work has been published for Tanzanian savannas, there are a couple of studies that used RapidEye and Landsat data in savannas, and these include Schmidt et al. (2012), Ramoelo et al. (2012), Naidoo et al. (2012), and Malatesta et al. (2013). However, most of these studies applied these datasets in environments different from the African savanna environments, and not specifically for REDD+ monitoring.

4.2.1 LiDAR Data

Airborne Light Detection and Ranging (LiDAR) technology is an active remote sensing technique capable of simultaneously mapping the Earth′s surface and overlying features, including vegetation and buildings, with sub-metre vertical accuracy. It uses a laser scanner that emits discrete pulses to the ground. From these, multiple returns are detected, which allows for the simultaneous mapping of the ground, vegetation, and other features on the surface. The distance is estimated based on the time between pulse emission and return detection (Lefsky et al., 1999; Bater and Coops, 2009).

LiDAR data has gained recognition in forestry and ecology as an effective tool for estimating a variety of vegetation metrics such as tree heights, biomass, crown size, leaf area index, and vertical canopy structure (Lefsky et al., 1999; Næsset and Økland, 2002; Ørka et al., 2009; Kim et al., 2009; Næsset, 2009; Wasser et al., 2013). The ability to estimate these forest attributes with a high level of accuracy has made it good for estimation of forest/non-forest. LiDAR provides useful

data for forest management, assessment and quantification, particularly at local scales (Rosette et al., 2009; Bunting et al., 2013b). LiDAR remote sensing for spatial distribution mapping of canopy characteristics has the potential of accurately and efficiently estimating tree dimensions and canopy structural attributes, whether at a local, regional or continental scale (Popescu et al., 2011; Wasser et al., 2013).

Data Acquisition

In 2012, an airborne LiDAR flight campaign was undertaken to acquire airborne laser scanning (ALS) data over Liwale. The campaign started on the 10^{th} February 2012 and was successfully accomplished on the 03^{rd} March 2012, almost a month-long. However, the actual data collection was done in only ten day that had to be spread across longer period due to persistent cloud cover over Liwale.

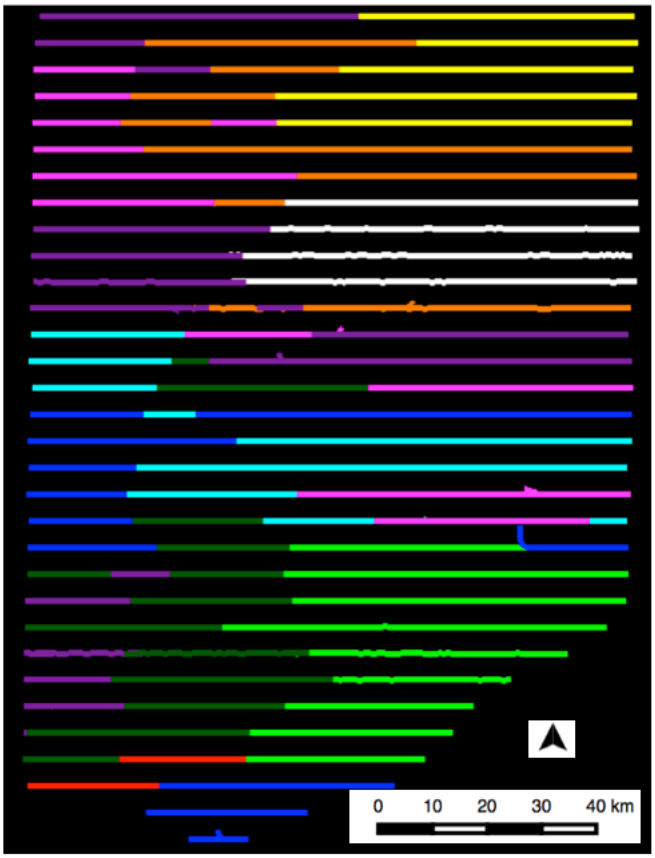


Figure 4.1: Coloured LiDAR trajectories, each colour showing a different flight date, to highlight the duration of the mission to persistent bad weather in Liwale.
Source: Terratec (2012)

Weather windows were relatively short, lasting for a duration of 1 to 2 hours, as depicted in Figure 4.1. LiDAR trajectories are shown in different colours, and each colour represent a different flight date (Terratec, 2012).

Since the LiDAR was only used for comparability purposes with products from the research, the LiDAR used coincided with RapidEye data availability over Liwale. Figure 4.1 is the spatial coverage of the 2012 LiDAR strips that was used for the comparison purposes in the research.

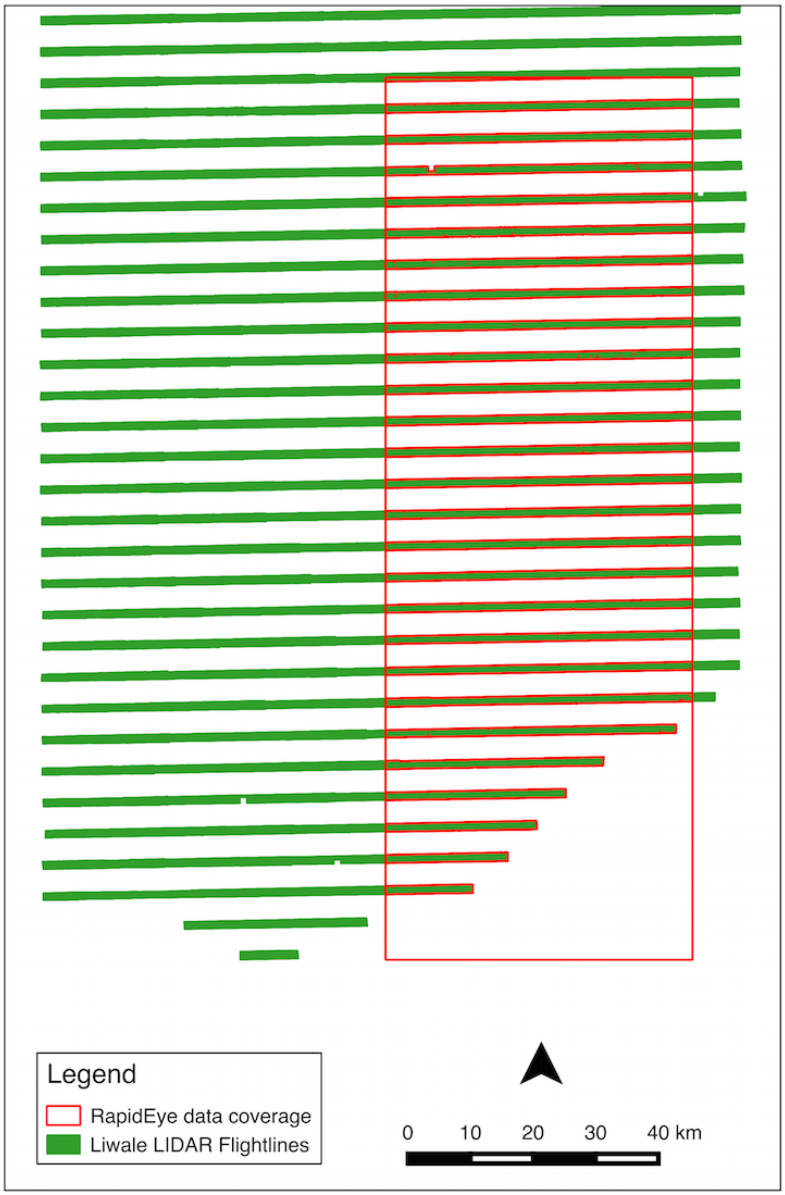


Figure 4.2: The 2012 LiDAR strips coinciding with RapidEye data over Liwale that was used in the study.

In 2014, another airborne LiDAR flight campaign was attempted over the same coverage area of Liwale. However, due to persistent cloud cover, and thus lack of good windows, the mission was aborted. Rather, only a small site within Liwale known as the SLU was successfully re-flown in August 2014. The SLU is made up of three LiDAR strips with the same width and coverage as the 2012 strips, and are shown in Figure 4.3. It is located in the south-east of the site of Liwale.

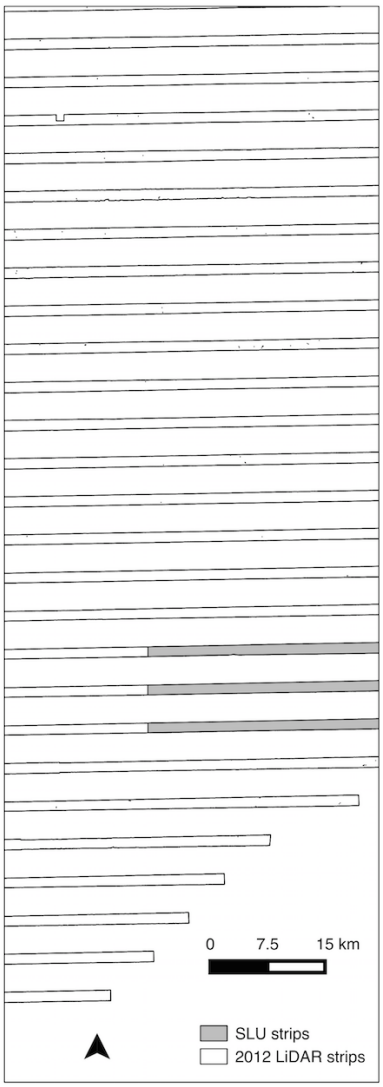


Figure 4.3: The SLU site in south-eastern Liwale for which LiDAR data was acquired in August 2014.

Sensor Specification and Calibration

Data was captured using an ALS70 laser scanner, serial number 7160, mounted on a twin engine Cessna 404 aircraft. Calibration parameters of the system were calculated by flying over a standard calibration area in Hamar, Norway. These parameters were also implemented in the ALS post processor for processing the raw dataset into Log ASCII Standard (LAS) format, which resulted in 1108 LAS data files (Terratec, 2012). Figure 4.4 shows the resultant LiDAR strips split into 3 km x 3 km blocks and each renamed into a unique code that first gives the strip number from top to bottom(001 to 032), followed by the block number on the strip from left to right.

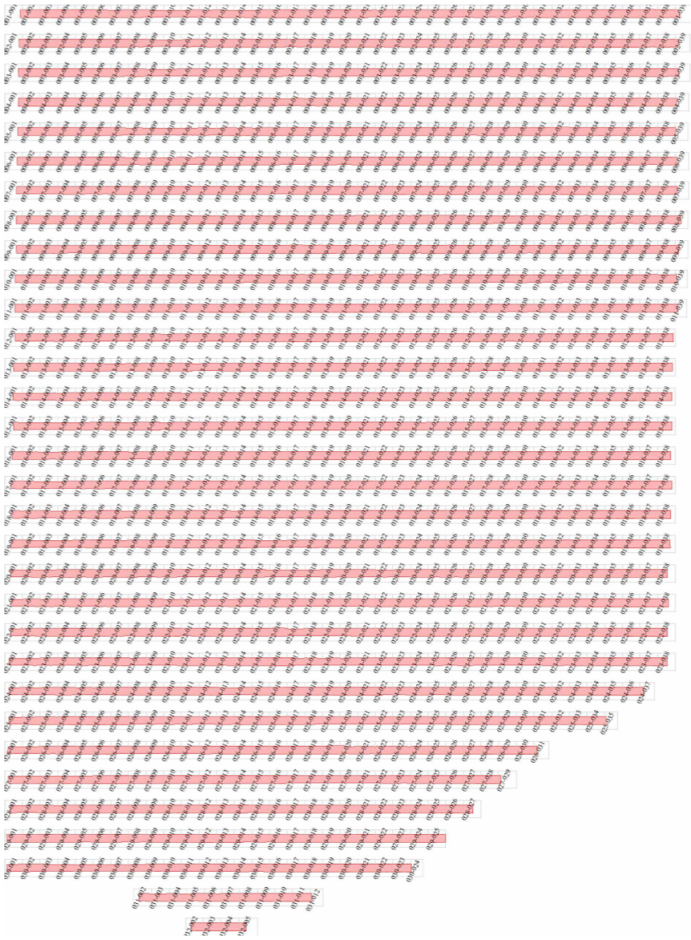


Figure 4.4: The 32 strips of LiDAR data acquired over Liwale in February 2012.
Source: Terratec (2012)

Data Preprocessing

The LAS data was preprocessed using Sorted Pulse Data Library (SPDLib), a set of open source software tools for processing laser scanning data (Bunting et al., 2013b). Its primary aim was to provide a set of tools supporting the processing of both airborne laser scanning (ALS) and terrestrial laser scanning (TLS) datasets for environmental remote sensing, particularly for vegetation metrics (Bunting et al., 2013a). One of the key features which differentiates SPDLib from other LiDAR software is its ability to process and store full waveform datasets alongside traditional discrete return data (Bunting et al., 2013b).

Being a large dataset of point densities, it was preprocessed in the HPC Wales high performance computing system (HPC Wales, 2015), and the SPDLib commands were perfomed on the system. LAS data were initially translated into individual unsorted pulse data (UPD), whereby the extent was calculated and pulses were built. In addition, the UTM 36S projection was assigned to the output UPD files, being the UTM zone Liwale lies within. The individual UPD files were then tiled using the number of rows (numofrows) and number of columns (numofcols) parameters, and finally gridding (spatially) each tile and outputting it in the sorted pulse data (SPD) format. The SPD format provides better storage and access to LiDAR data, as it stores the data within the hierarchical data format version 5.0 (HDF5) file format (Bunting et al., 2013a).

On the spatially sorted SPD files, the height field within both the pulse and point fields of each individual SPD file was then defined using the *spddefheight* command (Bunting et al., 2013a). The approach used was the Natural Neighbour interpolation, as demonstrated by (Bater and Coops, 2009), which interpolates a value for each point/pulse, and thus reducing artefacts as it generates a continuous surface. Even though the use of a Digital Terrain Model (DTM) of the same resolution as the SPD file bin size approach is the simplest, it was avoided because it can introduce artefacts resulting from the use of a series of spot heights.

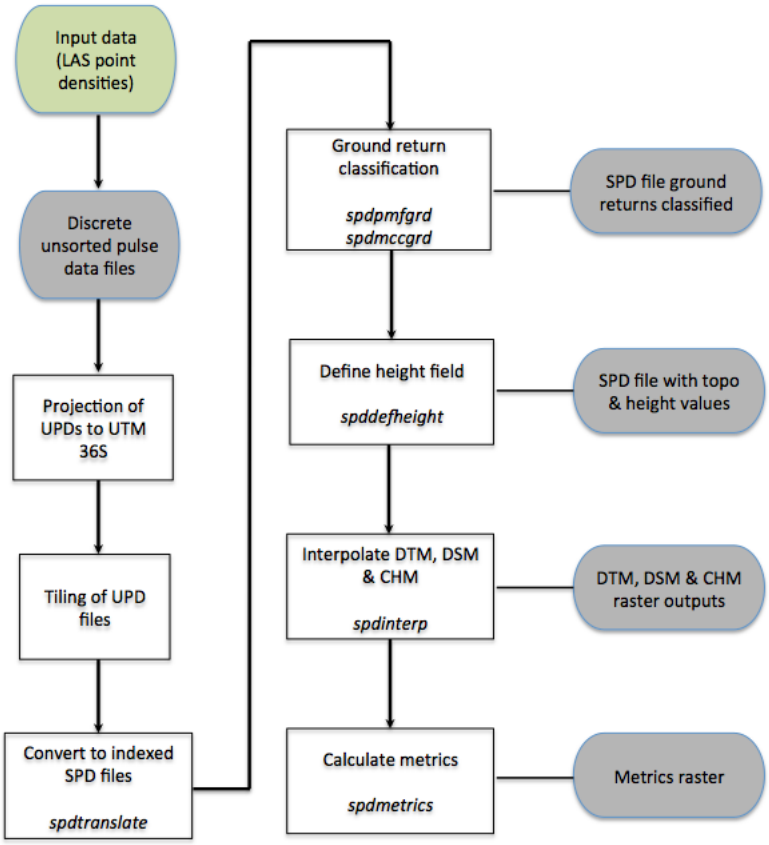


Figure 4.5: The processing chain followed during the preprocessing of the airborne LiDAR data in SPDLib.

The *spdinterp* command implemented in SPDLib was used to interpolate the raster surface from the classified ground returns and top surface points, to produce a canopy height model (CHM), digital terrain model (DTM) and digital surface model (DSM) for Liwale. These models were outputted as files in KEA format. The KEA file format was preferred over other image formats because it provides features of compression, ground control points, raster attribute tables, in-built image pyramids and support for large file sizes. Moreover, the KEA format can be easily integrated into GDAL as a driver (Bunting and Gillingham, 2013). Figure 4.5 illustrates the general processing chain that is undertaken during data preprocessing.

4.2.2 RapidEye Data

Properties of RapidEye Data

The RapidEye constellation is amongst the first of a new generation of operational small satellites that are being designed to meet the growing needs for purpose-driven remotely sensed satellite imagery (Gebhardt et al., 2012). It is made up of a constellation of five identical satellites operating in a sun-synchronous low earth orbit altitude of 630 km. Each satellite carries a 6.5 metres 5-band multi-spectral push broom imager that is capable of acquiring data at a wide range of roll angles (Dupuy et al., 2012). The five spectral bands it captures data in, as shown in Table 4.1, are blue, green, red, red-edge and near infrared (Förster et al., 2011; Ramoelo et al., 2012; RapidEye, 2012).

The RapidEye system is capable of daily access to any point on the surface of the Earth between latitudes range of -84^o and $+84^o$. This gives the sensor the ability to acquire more than 4,000,000 km^2 of imagery in a single day. Its capabilities therefore that include high-repeat, moderate-resolution and wide-area coverage make RapidEye imagery uniquely suited for a range of applications that requires high spatial resolution and temporal frequency data (Dupuy et al., 2012). RapidEye products are delivered with a pixel size of 5 m, either as level 1 (L1B) or level 3 (L3A or L3B). L1B data are delivered radiometrically calibrated and geometrically corrected, but not orthorectified whereas L3A and L3B are delivered having been both radiometrically calibrated and orthorectified. The difference between L3A and L3B is that the former is delivered as individual 25 km by 25 km tiles whereas the latter is large-scale product based on multiple images over an area of interest (RapidEye, 2012).

Table 4.1: The five spectral bands that make up RapidEye imagery, with a spatial resolution of 6.5 metres, but orthorectified to 5 metres.

Spectral band	**Band range**
Blue	440 - 550 nm
Green	520 - 590 nm
Red	630 - 685 nm
Red Edge	690 - 730 nm
Near infrared	760 - 850 nm

Source: RapidEye, 2012

For the complex open woodland environment of Tanzania, RapidEye was the appropriate sensor to trial its capability to adequately and accurately map and monitor these environments. Besides

the visible bands and the NIR band, RapidEye also acquires data in the red-edge range. The red-edge is a region in the red-NIR transition zone of vegetation reflectance spectrum and marks the boundary between absorption by chlorophyll in the red visible region, and scattering due to leaf internal structure in the NIR region. Greater sensitivity at higher chlorophyll content, detected as greener biomass has been observed for red-edge. Moreover, red edge position (REP) is influenced by chlorophyll content, leaf mesophyll structure, and leaf area index (LAI). Studies such as Sousa et al. (2012) and Kanke et al. (2012) reported that the integration of information from red edge increases classification accuracy. Sousa et al. (2012) specifically noted that different vegetation types appeared to be sensitive to the Red Edge channel and the derived indices.

Review of EyeFind Catalogue for Data over Liwale

During the review of available RapidEye data in the EyeFind catalogue, orthorectified L3A data was considered. Orthorectification is the process of removing the effects of image tilt and relief (terrain) effects, creating a planimetrically correct image, and thus features represented in their true positions (Slonecker et al., 2009). RapidEye is geometrically corrected using the globally available Shuttle Radar Terrain Mission (SRTM) image data and orbital information, and the target is to obtain accuracies in orthorectification of less than 1 pixel (RapidEye, 2012). Available L3A scenes for the globe can be viewed from the EyeFind catalogue, accessible from http://eyefind.RapidEye.com/. Through the use of a KML file of Liwale, scenes falling within the site were queried and reviewed, then the best scenes for each season were selected.

The approach initially adopted was to request quick-looks (Figure 4.6) from the Norwegian Space Agency and Kongsberg Satellite Services (project partners), who had acquired data over Liwale site. It was found that the earliest RapidEye data available for Liwale was acquired in 2010, then through the years into 2014. However, as the quick-looks in Figure 4.6 show, in each of the years, there was no total scene coverage. Since these forests of Liwale are predominantly open woodlands that are deciduous, the review aimed at producing a wet (leaf-on) and dry season (leaf-off) image for each year. The separation was to ensure that seasonality did not affect mosaicked scenes by giving variance within that mosaic for classes, especially the forest class.

Figure 4.6: Quicklooks of RapidEye imagery acquired over Liwale during different seasons: a) wet 2010, b) dry season 2010, c) wet season 2011 and d) dry season 2011, highlighting the shortage of appropriate RapidEye data in Liwale.

RapidEye Scenes for Liwale

The Liwale site (black polygon in Figure 4.7a) is covered by 24 tiles of L3A 25 km by 25 km scenes. After considering the distribution of available scenes (quick-looks) against the available financial resources for purchasing and acquiring data for the research, the initial area of interest was scaled down. The project partners were particularly interested in the south-western part (polygon marked by coral lines in Figure 4.7a) of the initial coverage area of interest. In this south-eastern part, there had been notable forest disturbance that had been observed in recent years. Furthermore, as part of the bigger project on MRV in Tanzania, there had been other datasets (e.g. Worldview and LiDAR) that had been acquired over that area. The target was to therefore source RapidEye data coinciding with these datasets spatially and, ideally, temporally. However, to adequately establish the use of optical data for forest mapping and monitoring, the selected study site also covered relatively undisturbed forest areas as well, and thus the decision to consider a larger area. This necessitated the inclusion of a greater number of scenes (polygon marked by purple lines in Figure 4.7a), with these numbered 1-19 tiles (Figure 4.7b).

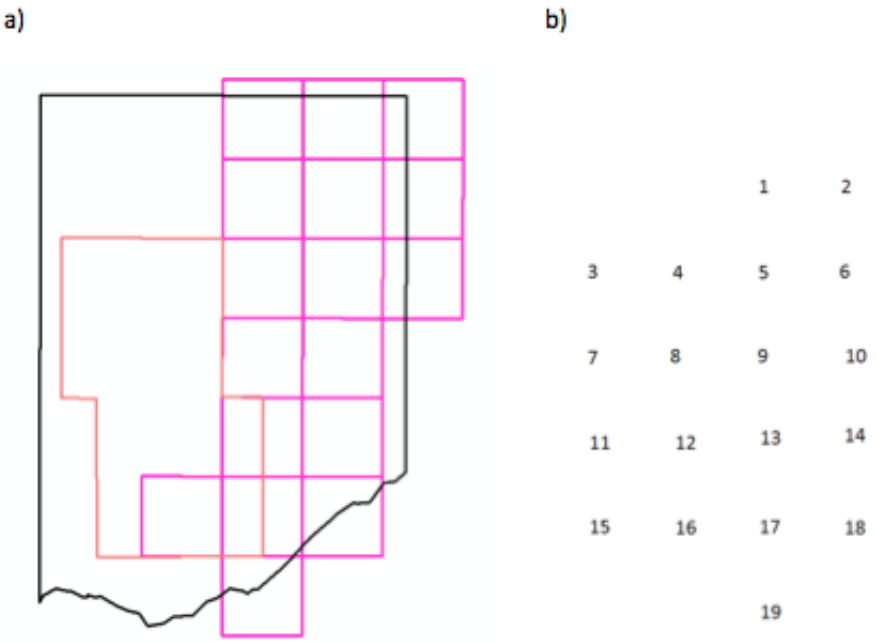


Figure 4.7: a) the 21 tiles covering the Liwale site, with b) 19 representing areas where disturbance has occurred and also a wide range of forest types exist.

Ideally, to account for both seasonal and phenological change within the tropical Liwale site,

monthly or at least quarterly mosaics per year are needed for these highly deciduous environments. However, as recommended by Huang et al. (2008), images acquired during leaf-off seasons should not be used in forest cover change analysis in areas where deciduous forests exist. Some tiles within the final site had a number of scenes on the EyeFind catalogue between the years 2010 and 2014. However, it was also noted that some tiles were without any scenes over the five years period. In addition, some of the tiles with scenes had too few or were all or mostly not suitable because of high cloud cover. The review revealed that cloud cover was a persistent challenge in the region both during the dry and wet seasons. Moreover, it was noted that available scenes were not evenly distributed across the year and there were several months without data. It was therefore, not possible to have monthly or quarterly mosaics over Liwale.

Another option explored was to generate seasonal mosaics. It was noted that successful acquisitions, or rather available scenes were mainly taken during the months of May, June, July, October, November, and December, with July and November being the most cloud free months. Liwale experiences the rainy season between November and May/June. The productivity of vegetation peaks around February, and remains high through May, and then sharply decline at the beginning of the dry season in June. Using this knowledge, scenes acquired between January and May each year were grouped for a wet season mosaic, and those acquired outside of these months were also grouped for a dry season mosaic. Thereafter, the best scene within each tile for the respective season in terms of cloud cover and missing data was chosen. Then mosaics by seasons for each year were constructed (Figure 4.8).

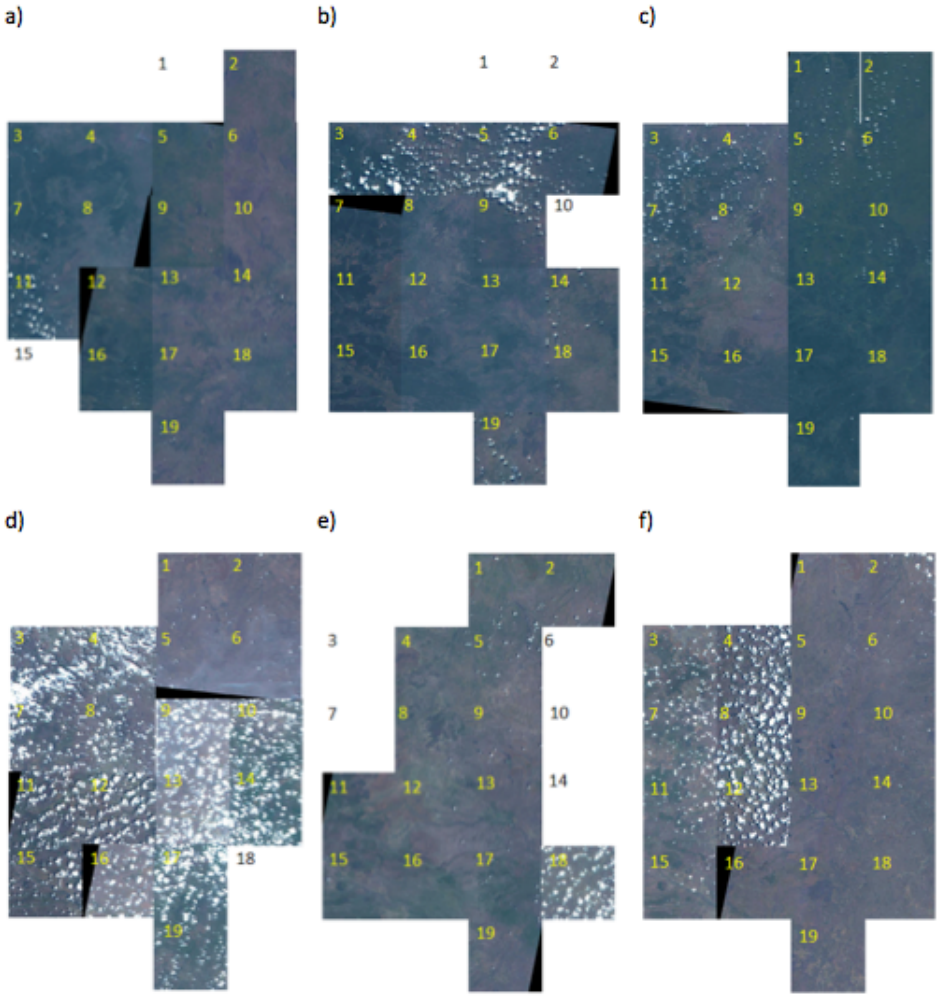


Figure 4.8: Best seasonal mosaic composites for Liwale for 2010-2012; a) wet season 2010, b) wet 2011, c) wet 2012, d) dry season 2010, e) dry season 2011 and f) dry season 2012.

Table 4.2 is a summary of dates the best scene was acquired per season for each tile over the years 2010-2014. The study area is made up of 19 tiles, and each year has its wet and dry season mosaic, meaning that we were supposed to acquire data for six mosaics, each made of 19 scenes. However, Figure 4.8 shows that some mosaics were missing data for some tiles. In total, without any missing data, there would have been 103 tiles acquired over the site, as summarised below in Table 4.2.

Table 4.2: Summary of the best L3A RapidEye scenes over Liwale for each wet and dry seasons of the years 2010-2014.

Liwale Tile ID	**EyeFind Tile ID**	**2010**		**2011**		**2012**		**2013**		**2014**	
		Dry season	*Wet season*	*Dry season*	*Wet season*	*Dry season*	*Wet season*	*Dry season*	*Wet season*	*Dry season*	*Wet season*
1	3734811	None	04th Nov	None	22nd Nov	18th May	23rd Oct	None	None	10th Aug	02nd Feb
2	3734812	09th July	04th Nov	None	22nd Nov	18th May	23rd Oct	None	None	22nd July	02nd Feb
3	3734709	10th July	28th Nov	05th July	None	06th Aug	11th Oct	13th Sept	14th May	16th Nov	14th May
4	3734710	10th July	28th Nov	03rd July	22nd Nov	06th Aug	17th Oct	13th Sept	15th April	10th Aug	14th May
5	3734711	20th June	04th Nov	03rd July	22nd Nov	18th May	23rd Oct	None	None	10th Aug	02nd Feb
6	3734712	09th July	04th Nov	03rd July	None	18th May	23rd Oct	None	None	22nd July	02nd Feb
7	3734609	10th July	28th Nov	07th July	None	06th Aug	30th Oct	13th Sept	21st May	03rd Aug	14th May
8	3734610	10th July	28th Nov	07th July	22nd Nov	06th Aug	17th Oct	24th June	30th May	10th Aug	14th May
9	3734611	09th July	21st Oct	09th July	22nd Nov	18th May	23rd Oct	None	None	10th Aug	02nd Feb
10	3734612	09th July	17th Dec	09th July	None	18th May	23rd Oct	None	None	22nd July	02nd Feb
11	3734509	10th June	24th Nov	07th July	22nd Nov	06th Aug	30th Oct	13th Sept	14th May	01st Aug	14th May
12	3734510	20th June	24th Nov	07th July	22nd Nov	06th Aug	09th Oct	24th June	None	10th Aug	14th May
13	3734511	09th July	21st Oct	07th July	22nd Nov	18th May	23rd Oct	13th Sept	30th May	10th Aug	02nd Feb
14	3734512	09th July	17th Dec	09th July	None	18th May	23rd Oct	None	None	10th Aug	02nd Feb
15	3734409	None	24th Nov	07th July	22nd Nov	06th Aug	30th Oct	13th Sept	30th May	03rd Aug	14th May
16	3734410	20th June	21st Oct	07th July	22nd Nov	06th Aug	23rd Oct	24th June	30th May	17th Nov	14th May
17	3734411	09th July	17th Dec	07th July	22nd Nov	18th May	23rd Oct	13th Sept	30th May	None	02nd Feb
18	3734412	09th July	None	09th July	03rd Oct	18th May	23rd Oct	None	None	None	02nd Feb
19	3734311	09th July	17th Dec	09th July	22nd Nov	18th May	23rd Oct	None	None	22nd July	None

Data Used

Due to seasonality in this environment where the forest is dominantly deciduous as well as the fact that LiDAR data had been acquired during the wet season of 2012 (February 2012), it was noted that the wet season 2012 dataset was the most suitable for developing a method for establishing a forest baseline for the area. However, since 8 of the tiles had the best scene acquired in August, they were found to have a great seasonal difference from the 11 that was acquired fairly within the wet season (May 2014). Eventually, for the purpose of forest estimation in Liwale, only 11 out of the 19 tiles making up Liwale had appropriate data that would enable the discrimination between forest and non-forest areas. The 11 tiles that were acquired on May 18^{th} 2014 were those with EyeFind tile IDs 3734311, 3734411, 3734412, 3734511, 3734512, 3734611, 3734612, 3734711, 3734712, 3734811 and 3734812. This is the eastern half of the whole site covered by the LiDAR data. The rest of the 8 tiles on the western half of the site were all acquired on August 06^{th} 2014, and it was found that most vegetation had shed the leaves by then. This made their use inappropriate as it would introduce seasonality.

For change detection, data covering the same area where a forest baseline had been established was needed. It was found that the 2013 wet season images that had been acquired on the 22^{nd} November had a lot of seasonal difference as the early rains had just come, and thus most trees were still without or with very minimal canopy cover. The 2014 wet season dataset, on the other hand, was acquired in February, during which vegetation cover is at its peak, and thus would have the same cover, or very close to the image acquired in May 2012. Therefore, the February 2014 dataset was used for developing a method for detecting change in the savanna forests of Liwale.

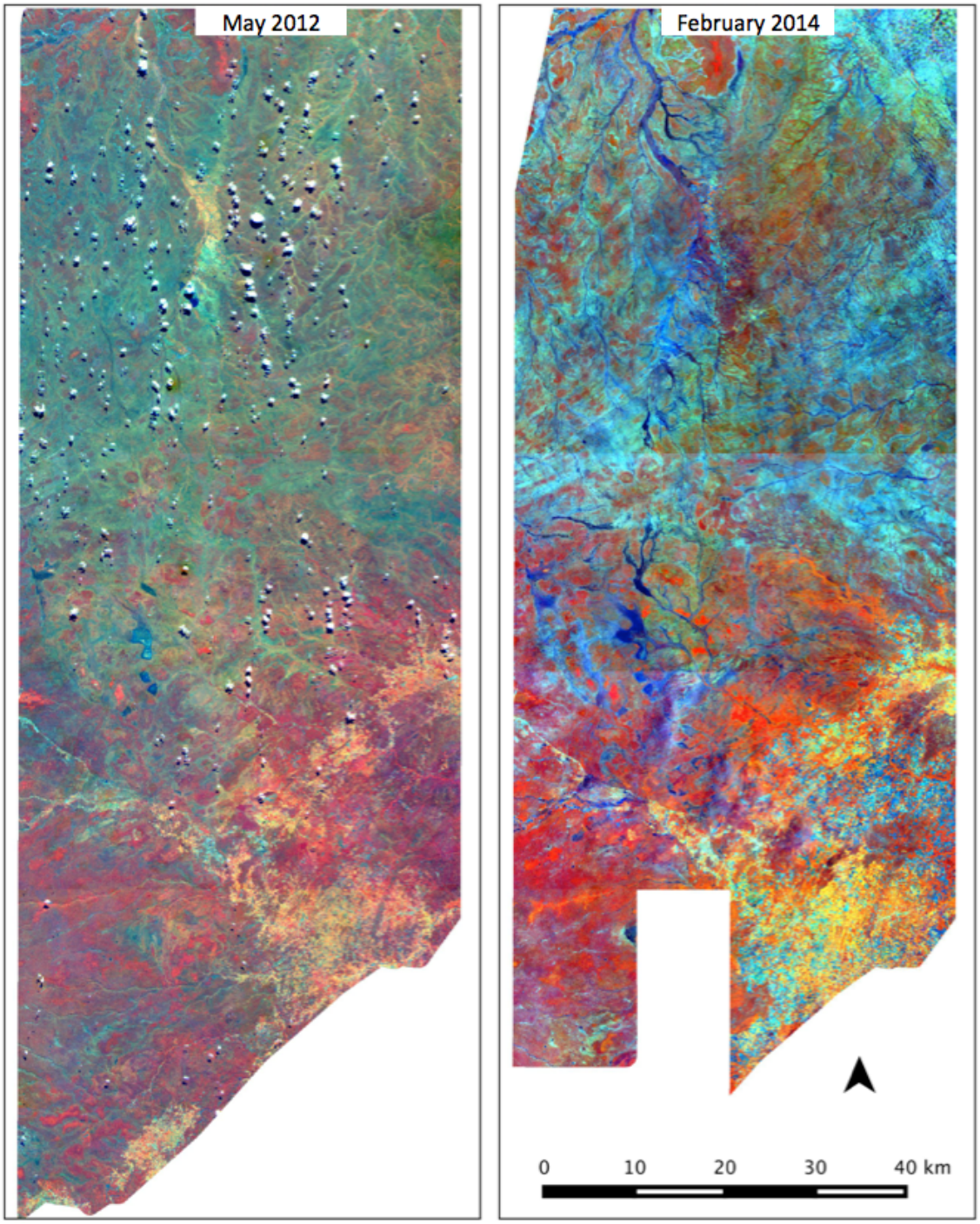


Figure 4.9: False colour mosaics of the RapidEye scenes for the years 2012 and 2014 that were used to establish the forest baseline, and to detect change respectively.

4.2.3 Landsat Data

The Landsat program is notably one of the longest running satellite data acquisition initiative that provides a comprehensive collection of global coverage of multispectral imagery. The program has provided continuous earth observation (EO) data for the past 41 years (Markham and Helder, 2012; Hansen and Loveland, 2012; Loveland and Dwyer, 2012). Data is systematically collected and archived following a global acquisition strategy (Banskota et al., 2014), dating back to 1972 (see Figure 4.10), even though there was a time when data was only available on demand than collected routinely.

The program that is run by the National Aeronautics and Space Administration (NASA) and the Department of Interior United States Geological Survey (USGS) (Irons et al., 2012; USGS, 2015b) has acquired millions of images over the decades of acquisitions, and these were made freely available since 2008 (Loveland and Dwyer, 2012; Wulder et al., 2012). Because of its global coverage and free availability, Landsat has become a key source of data for many countries. It is an invaluable resource with a suitable scale for landscape analysis, and thus suited for both local and global monitoring (Cunningham, 2014). Its continuity and spectral range has advantaged it as a great platform for forest mapping, and for both short-term and long-term forest monitoring (change). Markham and Helder (2012) noted that Landsat has been invaluable for characterizing and detecting changes in the land cover and land use globally.

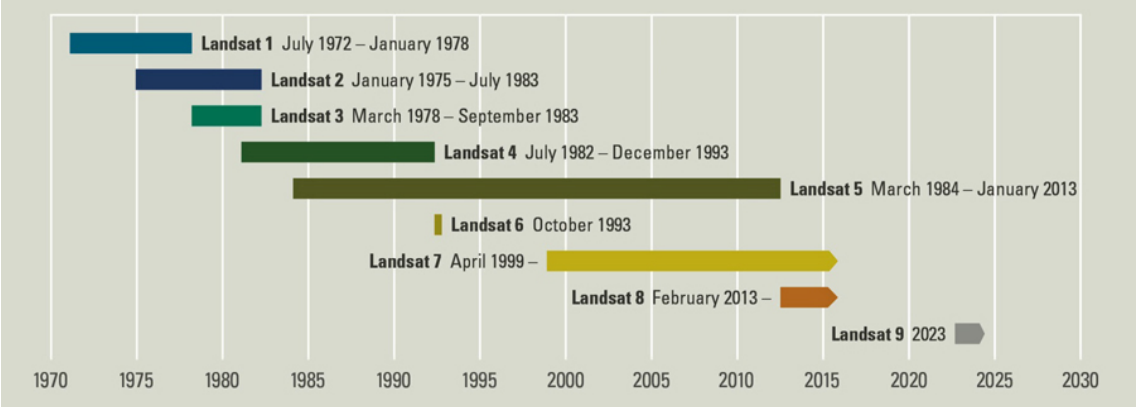


Figure 4.10: The different Landsat sensors and their operational timeline.
Source: USGS (2015b)

Properties of Landsat Data

Current and past Landsat missions have used four different sensor types for EO, summarized in Table 4.3. Far exceeding its 3-year design life, from 1972 through to the late 1992 and briefly in 2012 (USGS, 2014a,b), data was acquired using the Multi-Spectral Scanner (MSS), mounted on Landsat 1-5. For Landsat 1 and 2, its spatial resolution was approximately 80 m, with radiometric coverage in four spectral bands from the visible green to the near-infrared (NIR) wavelengths. Landsat 3 saw the introduction of an MSS sensor with a fifth band in the thermal infrared. Landsat 4 and Landsat 5 were launched in 1982 and 1984 respectively, with an addition of the Thematic Mapper (TM) to the MSS. The TM sensor collected data in the blue, green, red, NIR, SWIR1, SWIR2 and thermal regions of the spectrum. Its spatial resolution was 120 m for the thermal band and 30 m for the other six bands (Loveland and Dwyer, 2012; Banskota et al., 2014).

Landsat 7 introduced the Enhanced Thematic Mapper Plus (ETM+), launched in 1999. The ETM+ had similar spectral and spatial properties as TM, but had a 60 m spatial resolution thermal band and an additional 15 m panchromatic band. Even though Landsat 7 has also surpassed its 5-year design life and continues to acquire good quality data (Markham et al., 2004; Banskota et al., 2014), the Scan Line Corrector (SLC) of the ETM+ sensor has been inoperable since May 31^{st} 2003. Markham et al. (2004) noted that the SLC failure results in scenes with data gaps that range in size from two pixels at the scene centre to 14 pixels at the western and eastern scene edges. However, there has been methodologies such as Maxwell (2004), for filling SLC-off gaps through inference, using same scene spectral data.

The recent launch of the Landsat Data Continuity Mission (LDCM) in February 2013 provided an opportunity for a continuity in the use of these current and historic data for forest monitoring. Through a rigorous process of calibration and quality assurance (Markham et al., 2004), Landsat has ensured consistency in both sensor and image characteristics between Landsat 4 to Landsat 8 (Irons et al., 2012). This has resulted in fairly comparable Landsat observations from the TM, ETM+, and OLI sensors that enable time series analysis. Landsat 8 has onboard the Operational Land Imager (OLI) and thermal infrared sensor (TIRS). In addition to the bands on the ETM+ sensor, OLI collects data in a new short-wave infrared band (1.360 - 1.390 μm), located in a strong water vapour absorption feature, and thus play a big role in detecting cirrus clouds Banskota et al. (2014). Furthermore, widths of some OLI bands have been refined relative to ETM+ bands to avoid atmospheric absorption features (Irons et al., 2012). Another difference is that the data quality

(signal to noise ratio) and radiometric quantization (12-bits) of the OLI and TIRS instruments are higher than for TM and ETM+ (8-bit).

Table 4.3: A summary of the different Landsat sensors, showing their spectral bands, wavelengths in which they collect data, their pixel resolution, temporal resolution, and their operational timeline.

Sensors	**Bands**		**Wavelength (micrometres)**	**Resolution (metres)**	**Acquisition years**	**Revisit period**
Multispectral Scanner (MSS)	*Landsat 1-3*	*Landsat 4-5*			1972 - 2001	18
	Band 4	Band 1	0.5 - 0.6	60		
	Band 5	Band 2	0.6 - 0.7	60		
	Band 6	Band 3	0.7 - 0.8	60		
	Band 7	Band 4	0.8 - 1.1	60		
Thematic Mapper (TM)	*Landsat 4-5*				1984 - 1992	16
		Band 1	0.45 - 0.52	30		
		Band 2	0.52 - 0.60	30		
		Band 3	0.63 - 0.69	30		
		Band 4	0.76 - 0.90	30		
		Band 5	1.55 - 1.75	30		
		Band 6	10.40 - 12.50	120		
		Band 7	2.08 - 2.35	30		
Enhanced Thematic Mapper (ETM+)	*Landsat 7*				1999 - Ongoing	16
		Band 1	0.45 - 0.52	30		
		Band 2	0.52 - 0.60	30		
		Band 3	0.63 - 0.69	30		
		Band 4	0.76 - 0.90	30		
		Band 5	1.55 - 1.75	30		
		Band 6	10.40 - 12.50	60		
		Band 7	2.08 - 2.35	30		
		Band 8	0.52 - 0.90	15		
Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS)	*Landsat 8*				2013 - Ongoing	16
		Band 1	0.43 - 0.45	30		
		Band 2	0.45 - 0.51	30		
		Band 3	0.53 - 0.59	30		
		Band 4	0.64 - 0.67	30		
		Band 5	0.85 - 0.88	30		
		Band 6	1.5 - 1.75	30		
		Band 7	2.08 - 2.35	30		
		Band 8	0.52 - 0.90	15		
		Band 9	1.36 - 1.38	30		
		Band 10	10.60 - 11.19	100		
		Band 11	11.5 - 12.51	100		

Review of Existing Landsat Imagery over Tanzania

The Tanzanian environment has been reported to suffer from persistent cloud cover during both the wet and dry seasons in a year. Liwale site is almost completely covered by one Landsat scene; Row 166, Path 67. A thorough review of scene availability was performed for Liwale for the years 1990 - 2015. A review of the time series data for Liwale showed that it was not possible to have monthly images of the site throughout a year. The most viable approach was therefore to source a single date image for the wet season, taking into consideration canopy cover (seasonality). Table 4.4 is a summary of data availability by season (wet and dry seasons), and by sensor type (TM, ETM+ and OLI). As already highlighted, there is a distinct difference in vegetation canopy cover during the wet and dry seasons, and thus the need to separate them when performing the review. Scenes were classified into either available (appropriate) or unavailable (if with cloud cover greater than 30%). It was noted that there were relatively more scenes acquired during the dry season, and thus more availability during the season. But for forest studies, these scenes are inappropriate as vegetation is without canopy cover. However, during the wet season, there were relatively less acquisitions, especially with the TM and ETM+ sensors in earlier years.

The introduction of OLI sensor saw a hike in scene acquisition during the wet season, as it was noted for 2014 and 2015. However, cloud cover, and ultimately availability is still a challenge. Even though a total of 9 scenes were acquired in 2014, only 2 would be usable, and out of 11 in 2015, none had less than 50% cloud cover. For that range of over 20 years, the maximum number of available scenes from the TM sensor was never greater than two, and for 12 years no scene was available. Landsat 7 seemed to provide relatively at least a single scene per wet season, which unfortunately is not adequate post 2003 when the SLC failed as data gaps need to be filled up before use.

Table 4.4: A summary of all available Landsat scenes for Liwale between the years 1990 and 2015, grouped into either wet or dry season.

Year	Wet season						Dry Season					
	No. of available images			*No. of cloudy images*			*No. of available images*			*No. of cloudy images*		
	TM	*ETM+*	*OLI*	*TM*	*ETM+*	*OLI*	*TM*	*ETM+*	*OLI*	*TM*	*ETM+*	*OLI*
2015	-	1	0	-	4	11	-	1	1	-	0	0
2014	-	1	2	-	8	7	-	7	9	-	7	4
2013	0	2	-	1	6	-	7	7	-	6	5	-
2012	0	1	-	0	5	-	0	9	-	0	1	-
2011	0	1	-	1	3	-	1	3	-	0	3	-
2010	1	3	-	3	2	-	1	2	-	2	1	-
2009	1	2	-	0	5	-	4	6	-	2	4	-
2008	2	4	-	6	3	-	6	3	-	2	4	-
2007	1	2	-	4	2	-	1	3	-	1	4	-
2006	0	1	-	4	2	-	7	4	-	5	3	-
2005	0	0	-	2	8	-	3	5	-	1	1	-
2004	1	2	-	2	2	-	2	3	-	2	3	-
2003	0	1	-	0	2	-	1	2	-	2	1	-
2002	0	2	-	0	7	-	0	4	-	0	7	-
2001	0	1	-	0	5	-	0	4	-	0	7	-
2000	0	2	-	0	5	-	1	3	-	1	3	-
1999	0	-	-	3	-	-	0	2	-	0	2	-
1998	2	-	-	3	-	-	5	-	-	3	-	-
1997	1	-	-	5	-	-	4	-	-	1	-	-
1996	0	-	-	2	-	-	0	-	-	2	-	-
1995	2	-	-	6	-	-	3	-	-	0	-	-
1994	1	-	-	2	-	-	5	-	-	3	-	-
1993	1	-	-	0	-	-	5	-	-	0	-	-
1992	1	-	-	0	-	-	1	-	-	0	-	-
1991	2	-	-	2	-	-	2	-	-	1	-	-
1990	0	-	-	1	-	-	3	-	-	0	-	-

** Grayed out cells are images acquired with ETM+ sensor after the SLC stopped working

Landsat Data Used

Since the aim with Landsat was to further trial the methods and techniques developed using RapidEye data for both forest mapping and change detection, the most ideal situation would have been to use Landsat data coinciding with the 2012 and 2014 RapidEye wet season data. The idea was therefore, to have a 2012 Landsat wet season image coinciding with the 2012 RapidEye wet season image that had been used to establish a baseline for Liwale. Similarly, to trial scaling out the method for detecting forest change (disturbance), a 2014 Landsat wet season image would have been ideal to compare the change product with that produced from a 2014 RapidEye wet season image. However, data availability in Liwale was found to be a major issue, first due to cloud cover, and secondly, due to seasonality.

Table 4.4 shows that there was only one scene acquired using ETM+ available, which was not

appropriate for use on its own as gap-filling would not be possible. The other five scenes from the ETM+ sensor were not usable due to very high cloud cover of greater than 30% of the whole scene. 2013 did not have any TM scenes for the wet season too, and the two available ones for ETM+ could not be gap-filled to produce a filled scene because they were too far apart, and had seasonal difference, which would likely introduce error during the forest/non-forest classification/mapping.

With the launch of Landsat 8 early 2014, there was an improvement on scene availability over Liwale, which resulted in 9 scenes being acquired and made available through USGS Earth Explorer website (USGS, 2015a). However, to underpin the issue of cloud persistence in the region, only two of those had less than 30% cloud cover. From the two, an image acquired on the 12^{th} May 2014 was found to be the most appropriate as it had less cloud cover and without seasonality. Having gone with a 2014 image for trialling the establishment of a forest baseline, 2015 was considered for trialling the change detection method. However, Figure 4.11 shows that the whole of the 2015 wet season images had more than 50% cloud cover, which made them inappropriate for the task.

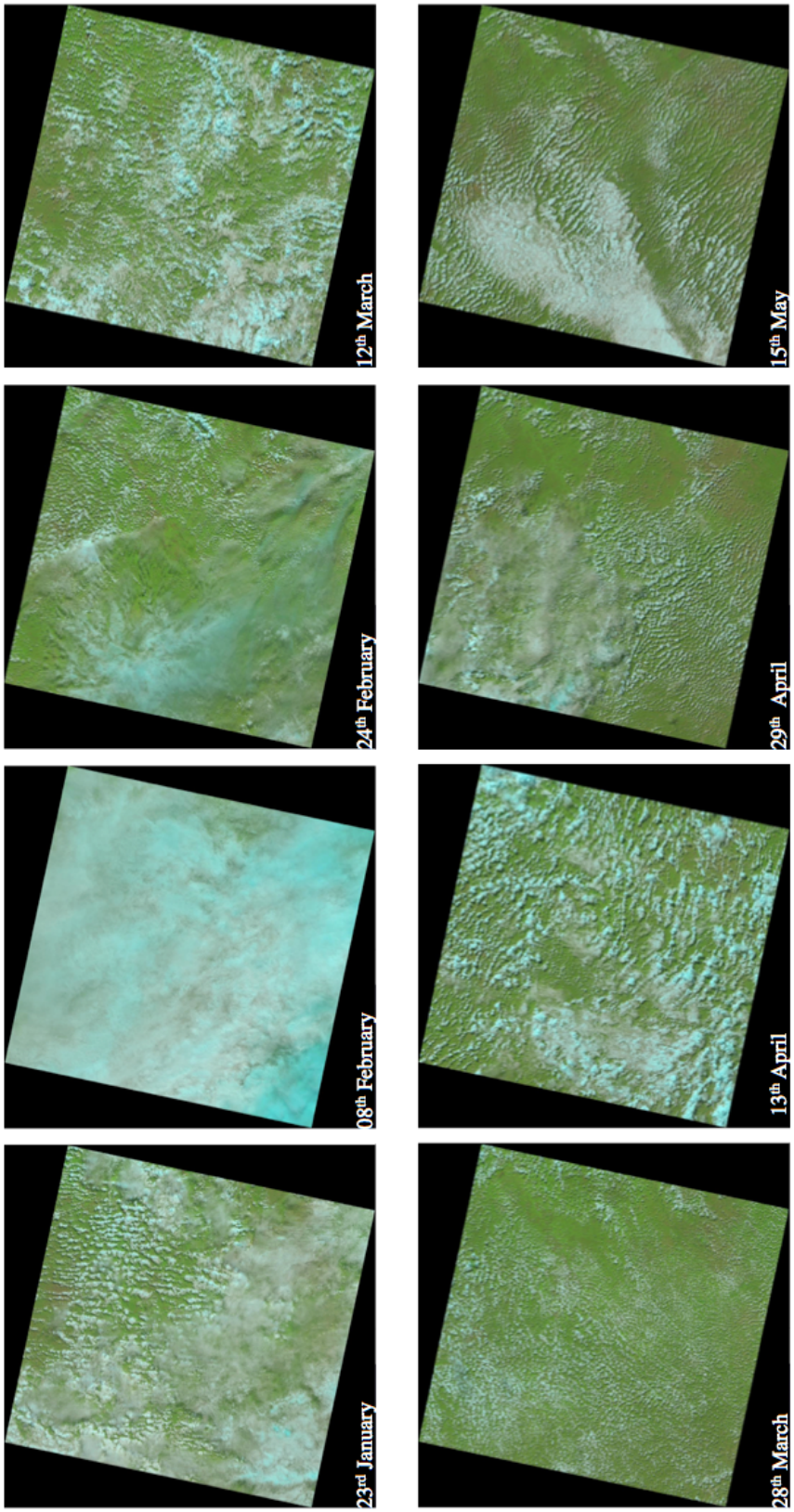


Figure 4.11: A collection of all the Landsat 8 images of the 2015 wet season period, showing that all the scenes had more than 50% cloud cover.

Of these images, the one that seemed to have the least cloud cover is the one that was acquired on the 28^{th} March 2015. However, Figure 4.12 shows that this image was inappropriate too as the cloud cover, even though light, was spread across the scene. Masking out clouds and shadow would result in about 50% of data lost.

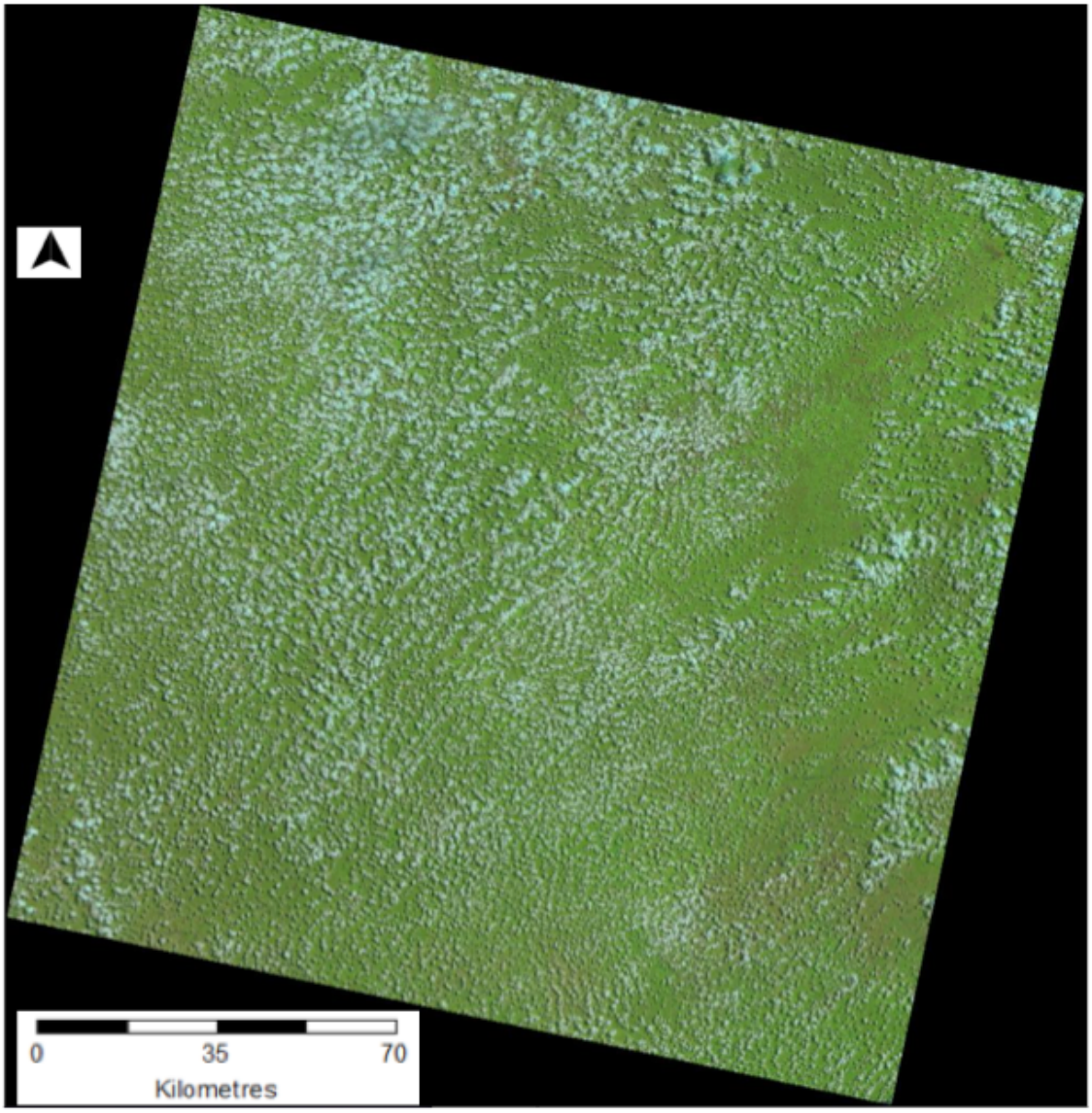


Figure 4.12: Landsat image acquired on the 28^{th} March 2015, showing that even though it looks to be the best from the 2015 collection, the amount of cloud cover (>50%) makes it inappropriate for use.

This led to modifying the approach of using a single date image to establish a forest baseline, and then use the forest baseline and an image of a later data to detect change, into using an earlier

image to detect change down the years. Even though good, and sometimes cloud-free wet season images were found over Liwale in earlier years, due to variance in weather conditions between years, it was found that scenes within the wet season had some seasonal difference. This would introduce error as some forested areas would be picked up as change due to seasonality. The best scene that was not affected by seasonality was found in the year 1998, acquired on the 16^{th} May 1998. Figure 4.13 shows the two images that were eventually used to establish a forest baseline and for change detection respectively.

Table 4.5 is the summary of the datasets that were used in the study, showing when each RapidEye and Landsat imagery used for either the establishment of the forest baseline or the detection of change, were acquired. The table further shows when the LiDAR data, whose CHM products were used to compare against the forest baseline and change products, were acquired too. Lastly, the table shows when field data that was handy for developing rulesets for undertaking the classifications was undertaken.

Table 4.5: A summary of the different datasets that were used in the study, and their time of acquisition.

Data Use	**Data**			
	RapidEye	***Landsat***	***LiDAR***	***Fieldwork***
Forest baseline establishment	18th May 2012	12th May 2014	February 2012	February 2013
Change detection	02nd February 2014	16th May 1998	August 2014	

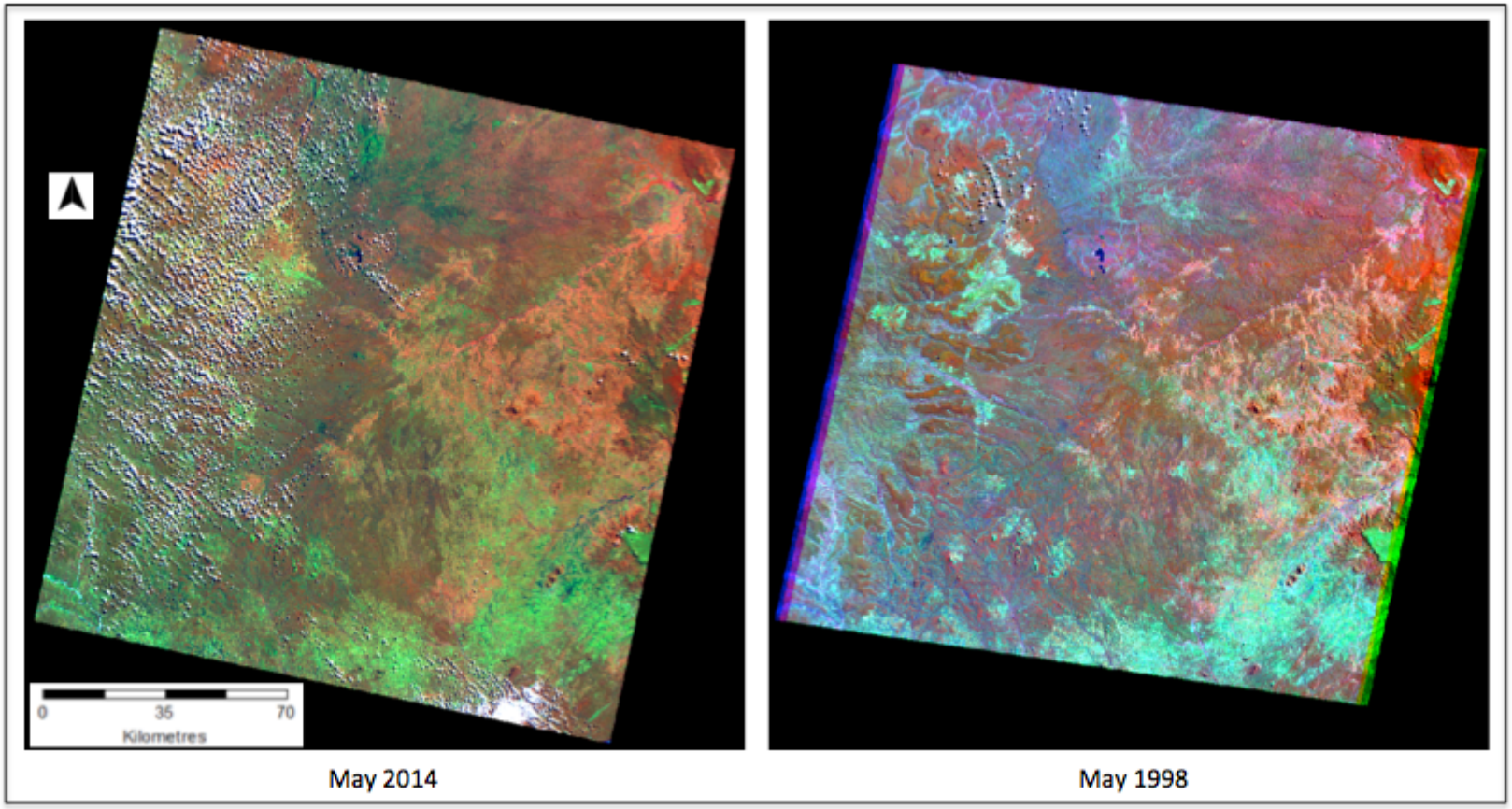


Figure 4.13: The May 2014 Landsat image that was used for establishing a forest baseline for Liwale, and a May 1998 Landsat image that was used with the established forest baseline for change detection.

4.3 Preprocessing of RapidEye and Landsat Scenes

4.3.1 Introduction

Image reflectance, that is, the spectral reflectance signals of different features within an image, are influenced by a number of different factors during data acquisition. These reflectance signals, whether insolation or outgoing radiation, are influenced at the top of the atmosphere, atmosphere and land surface. Among the factors that influence it are topography, slope, cloud cover, atmospheric water content, solar elevation angle, and the sensor itself (Afify, 2011; Alqurashi and Kumar, 2013). This, therefore, has an effect on the radiation that reaches the surface and that which is eventually received by the sensor. It necessitates the need for correction of the remote sensing data before it can be used. Norjamäki and Tokola (2007) state that the accuracy of satellite image-based environmental analysis, be it LULC mapping or forest inventory and monitoring, depends on the quality of the satellite data interpreted. This lines up with the observation by Song et al. (2001), Magdon et al. (2014) and Banskota et al. (2014) that remote sensing data correction has become indispensable and a crucial part of image processing.

For this research, both RapidEye and Landsat scenes used were first preprocessed and corrected for Top of Atmospheric (TOA) effect, then to Surface Reflectance (SREF), using Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) software. This is an open source software package developed by Aberystwyth University (Bunting et al., 2014a). For both Landsat and RapidEye satellite data, rigorous tests were undertaken across different environments in different continents (Bunting et al., 2014b; Bunting, 2015). This study therefore employs 6S implemented in ARCSI to preprocess both RapidEye and Landsat data of Liwale.

4.3.2 Conversion to Radiance

Both RapidEye and Landsat data are provided as digital numbers (DNs) images (RapidEye, 2012; Banskota et al., 2014), but stored differently as the sensor sensitivity range and manufacturer design varies from sensor to sensor (Cunningham, 2014). DNs enable providers to scale the data to reduce its size and thus avail more storage space. The first step undertaken was to calibrate the data from DNs to radiance. Radiance is defined as the amount of energy received by the sensor per second (W) per steradian (sr) per square metre (m^2), with the units $Wsr^{-1}m^{-2}nm^{-1}$.

RapidEye Conversion

RapidEye data are quantized at 12 bit and stored with a bit depth of up to 12 bit. Then on-ground processing is undertaken to perform radiometric corrections, which results in images being scaled to 16 bit. This step, according to (RapidEye, 2012), converts pixel DNs derived directly from the sensor into values directly related to absolute at sensor radiances. To convert the DN of a pixel to radiance, the DN value is multiplied by the radiometric scale factor, as follows:

$$RAD(i) = DN(i) \times radiometricScaleFactor(i) \tag{4.1}$$

where

- radiometricScaleFactor(i) = 0.01

Landsat Conversion

Landsat on the other hand, is converted from DN to radiance using the linear equation given as:

$$L = (\mathrm{DN} \times G) + I. \tag{4.2}$$

Where:

- L = Radiance
- G = Gain
- I = Intercept

Worth noting is that the values of the gain (G) and intercept (I) differs for each band, and these are provided in the image header file.

To convert each individual band into radiance, an ARCSI command was executed, outputting the converted images to the KEA file format, using the calibration factors specified in the image file header.

4.3.3 Conversion to Surface Reflectance

Background

The electromagnetic radiation (EMR) signals collected by satellites in the solar spectrum are modified by scattering and absorption by gases and aerosols while travelling through the atmosphere from the Earth's surface to the sensor (Themistocleous et al., 2012). Radiometric correction is a prerequisite for generating high-quality scientific data. It corrects for any distortion or degradation stemming from the characteristics of the imaging system (sensor) and imaging conditions (Pons et al., 2014). This makes it possible to discriminate between the data product's artefacts and real changes in the Earth surface, and therefore ensure the production of accurate land cover maps and change detection. Prior to data analysis, initial processing on the raw data is usually carried out. This is commonly known as image preprocessing, and Lillesand et al. (2008) terms it image rectification and restoration. It involves correction for geometric distortion, data radiometric calibration, and noise elimination in the data (Hadjimitsis et al., 2010).

The standard procedure for satellite data pre-processing requires, first the calibration to radiance (W m^{-2} sr^{-1} m), followed by either conversion to top-of-atmosphere (TOA), relative or surface reflectance (SREF) (Bunting et al., 2014b). For conversion to surface reflectance, atmospheric correction typically involves a radiative transfer (RT) model. Commonly used models include 6S, ACORN, ATCOR, Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS), LOW-resolution atmospheric TRANsmission (LOWTRAN/MODTRAN), and Quick Atmospheric Correction (QUAC). The parameterisation process is based on either standard or actual atmospheric properties that are sourced from referencing meteorological or other direct or indirect satellite measurements of atmospheric constituents.

Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI)

To correct an image to surface reflectance, an atmospheric model can be used to account for the effects of scattering and absorption in the atmosphere. This study used Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) software, to parametise the 6S radiative transfer model. ARCSI software provides a command line tool for the atmospheric correction of Earth Observation imagery. The aim of ARCSI is to provide as automatic as possible method of

retrieving the atmospheric correction parameters and using them to parameterise 6S.

Bunting (2015) developed the ARCSI software that implements 6S to pre-process remote sensing data to surface reflectance. For both Landsat and RapidEye satellite data, rigorous tests were undertaken, and different preprocessing techniques developed. This study therefore employs 6S implemented in ARCSI to preprocess both RapidEye and Landsat data of Liwale. ARCSI uses the 6S radiative transfer source code that is freely available for download at http://6s.ltdri.org, to correct for surface reflectance.

6S Surface Reflectance Correction

Numerous studies advocate for scene-based parametisation than reliance on ground data (Tripathi et al., 2005; Themistocleous et al., 2012; Pons et al., 2014; Bunting et al., 2014b). This approach corrects for atmospheric effect through the use of algorithms utilizing information sourced from the image itself (Norjamäki and Tokola, 2007). A study by Song et al. (2001) that compared different algorithms for correcting atmospheric effects revealed that more complicated algorithms did not necessarily lead to improved performance of classification and change detection. Rather it recommended simple dark object subtraction (DOS), with or without the rayleigh atmosphere correction, or relative atmospheric correction for classification and change detection applications.

This is the algorithm that was implemented in ARCSI through 6S in order to correct both RapidEye and Landsat images to surface reflectance. Moran et al. (1994) and Lei et al. (1998) concurred with the recommendation as they found that the effect on accuracy of atmospheric correction obtained from a standard atmosphere model was comparable to that obtained from an in-situ model.

A number of parameters are required for 6S, and these are; atmospheric profile, aerosol profile and Aerosol Optical Depth (AOD), ground reflectance, geometry, altitude, and wavelength. Most of these parameters are calculated within ARCSI or derived from the header file (e.g., geometry, altitude and wavelength). There are a number of preset options for atmospheric profile, aerosol profile and and ground reflectance in ARCSI. Figure 4.14 is a flowchart of the steps undertaken during the preprocessing of both RapidEye and Landsat scenes used in the study.

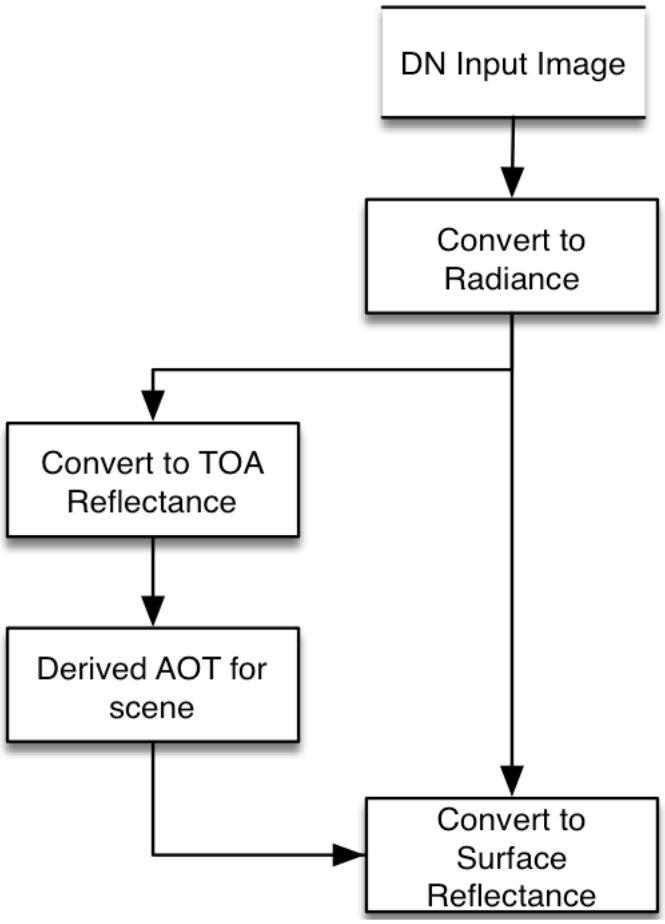


Figure 4.14: The processing chain that was followed during the preprocessing of both RapidEye and Landsat scenes in ARSCSI software.

4.3.4 Testing the Quality of ARCSI-derived Products

The rationale for the choice of ARCSI as the preprocessing tool is that it is a scriptable and flexible software, that was found to produce good results for both RapidEye and Landsat when tested for sites in Tanzania and other parts of the globe (Bunting et al., 2014b).

Worth noting is that there was no available spectral field data for Liwale or Tanzania to use in testing the quality of the corrected RapidEye and Landsat products, which would have been done through comparing the field spectra with the ARCSI-derived spectra. Therefore, the ARCSI results were first visually assessed to make sure the values and shapes of spectral curves were as expected and published, an example of such visual comparison is Figure 4.15, and the results lined up with

those of (Bunting et al., 2014b). The figure shows an example of the spectral curves of a pixel within a forest area, for top of atmospheric correction and surface reflectance from a May 2012 RapidEye image.

After the visual assessment, an approach that would have ideally been used would be the selection of invariant features from time-series data and the stability of the reflectance over time assessed. As shown in the review, RapidEye has a poor time-series data over Tanzania, and the Landsat time series is greatly affected by seasonality of the environment. An alternative approach of a direct pixel-to-pixel comparison of Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al., 2013) and ARCSI-derived surface reflectance values was used. However, since there was lack of temporal data or other surface reflectance products to compare against, this comparison was for Landsat scenes only. Since the same methods were used for both sensors, it was assumed that if they worked well for Landsat, they would also work for RapidEye to the same level of accuracy.

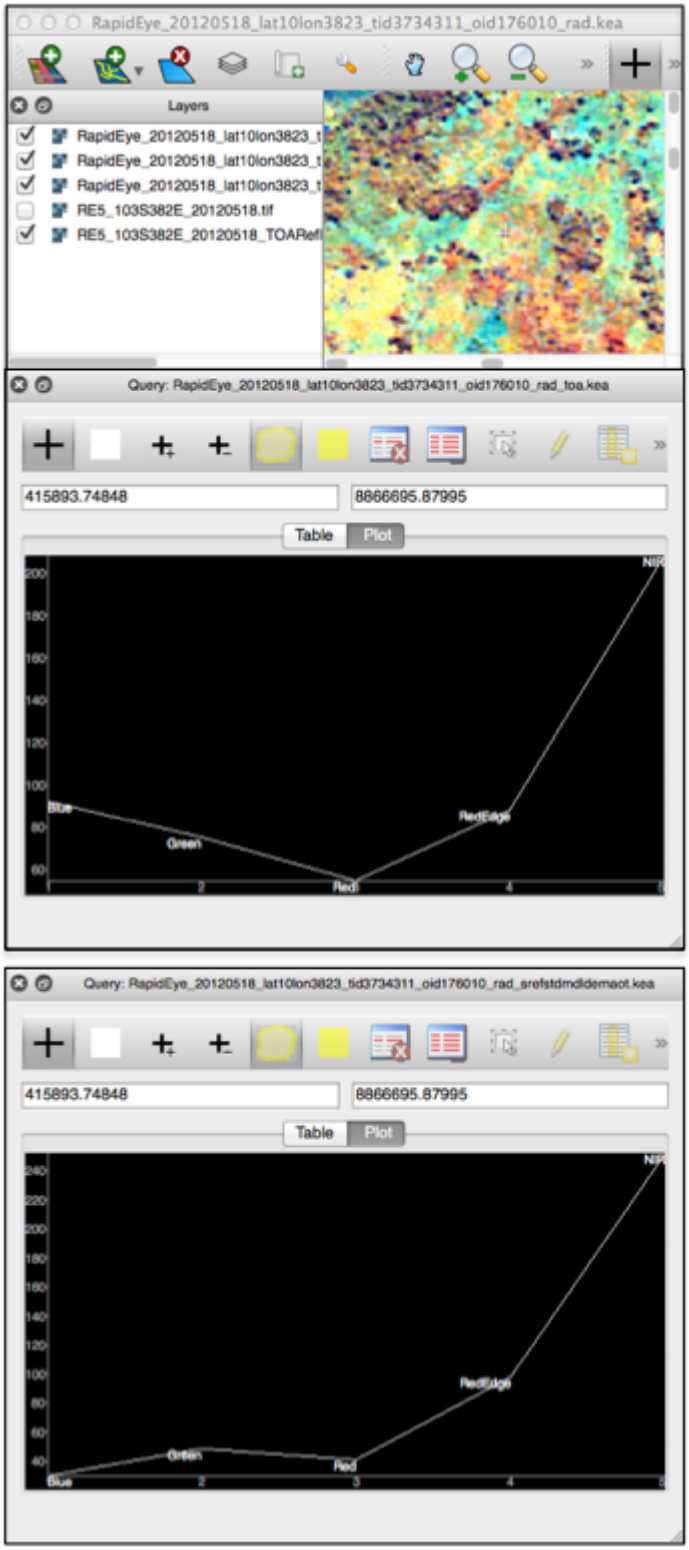


Figure 4.15: A comparison of ARCSI-derived spectra of a RapidEye data product that was corrected to top of atmosphere (middle) and surface reflectance (bottom).

4.4 Fieldwork

4.4.1 Introduction

In the month of February 2014, a fieldwork was undertaken in Liwale, in cooperation and partnership with the NAFORMA team made of experts from Sokoine University of Agriculture, Tanzania Forestry Research Institute (TAFORI), and field experts from Tanzanian government ministries. This was during a NAFORMA field campaign, and it provided a more in depth view of the study region, and a field dataset was collected.

4.4.2 Fieldwork Approach and Data Collection

During the fieldwork, a Garmin GPS with an accuracy of ±3 m accuracy was used to capture 435 waypoints within the Liwale study site, which covers approximately 16,000 km^2. The waypoints were mainly of centres of plots of a diameter of about 30m of different land cover and land use types and practices in the area, ranging from newly cleared forest areas to intact forest, wetlands and agricultural fields (*shambas*). Points of interest such as individual trees, felled trees, burnt trees and tree stumps were also marked. In 133 of the waypoints, a recording sheet was used to classify for the LCCS.

4.4.3 Fieldwork Results

Forest Types in Liwale

Without the influence of human activities and wildfires, Liwale forest type is predominantly miombo woodland vegetation type. Figure 4.16 shows some of the common vegetation covers in relatively less disturbed Liwale environment. The woodlands vary across the landscape, with some being sparsely populated woodlands (Figure 4.16b) while at other areas fairly closed woodlands (Figure 4.16c). More open woodland areas were noted mostly on the northern half of the study site, towards and within Selous Game Reserve. There is also a lot of open grassland with very sparse vegetation and shrubs (Figure 4.16a). In this part of the study site, there is minimal human disturbance, with lesser human population settled there.

Figure 4.16: Forest types in near pristine parts of Liwale where there has been minimal human activity.

The most common forest disturbances in the northern part of Liwale site include log harvesting, wildlife, and wild fires. Fallen and dead trees as a result of being toppled by herbivory animals such as elephants or being burnt by wild fires were observed (Figure 4.17 and Figure 4.18). However, the scale of the disturbance is relatively low compared to the central and southern part of the Liwale site, which has a higher development, villages, and thus population settled there. It was observed that most of the villages in and around the study site practised agricultural activities, especially shifting cultivation. This was most pronounced in the southern part of the site. This has led to a notable amount of forest disturbance in the area, both from deforestation and degradation, which is for infrastructural development and subsistence farming. The woodlands are deciduous in nature; shedding their leaves during the dry season lasting from June to late October or early November. The southern part of Liwale site is predominantly closed forest with a lot of understory (Figure 4.16d). This part of the site is still deciduous, but there are parts of it with semi-deciduous forest clusters.

Figure 4.17: An example of a burnt down tree which highlight the scale of some forest fires that burn the landscape.

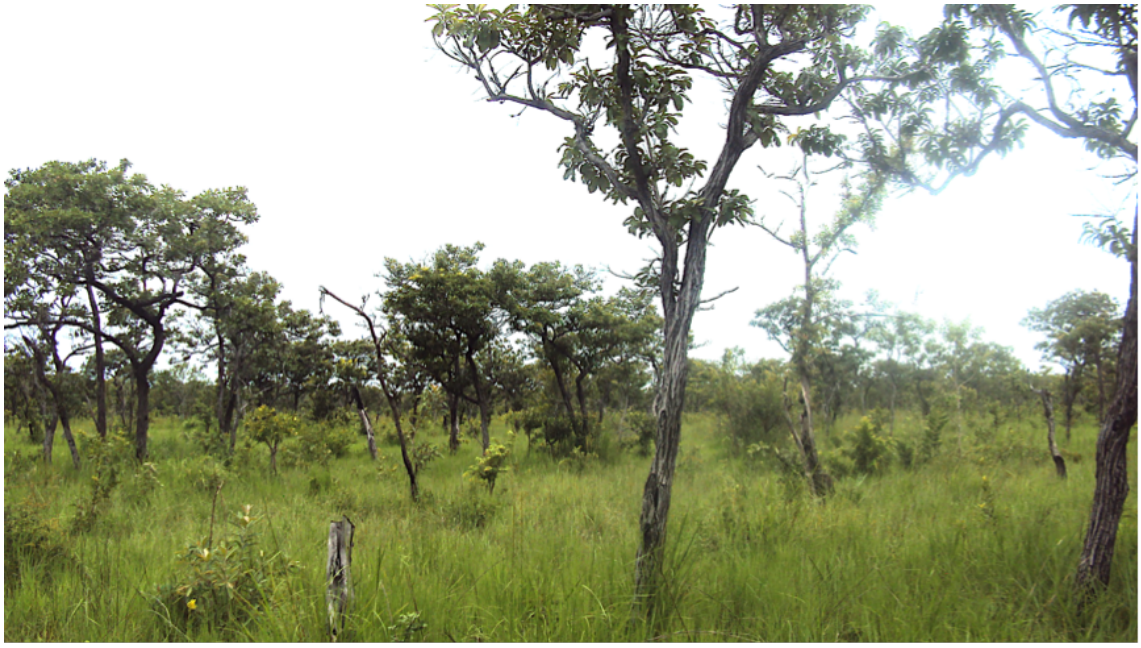


Figure 4.18: A tree stump is what remains of a tree that was burnt to ashes, even dried up stumps and charred tree stems in the background show the effects of wild fires on woodlands in the northern parts of Liwale site.

Common Practices Encouraging Forest Degradation

The primary form of degradation in Liwale is subsistence agriculture, discussed in Chapter 3. To date, most subsistence farmers in rural areas in Tanzania practice shifting cultivation. Lands ranging up to 2 hectares are cleared into agricultural fields, commonly known as *shambas* in the

native language. The most commonly cultivated crop is maize, even though farmers tend to practice intercropping. Due to lack of scientific data on forest degradation, it's true scale and dynamics are still not known or fully understood. In rural areas like Liwale where subsistence agriculture is the main activity responsible for forest disturbance, deforestation and degradation are somehow intercrossing, whereby both forms of disturbance are taking place side by side or concurrently. Their main causes include clearing for agriculture, overgrazing, wildfires, mining and wood extraction for charcoal, timber, poles and firewood (Zahabu, 2008). Of these causes, Mwampamba (2009) noted that shifting cultivation alone contribute more than 50% of the deforestation in Tanzania.

Generally, shifting cultivation is dynamic, taking different forms. At national level, the forms are dependent on location, crop species diversity, crop type, field size, cultural pattern, population density, socio-economic pressure and methods of field preparation (Mwampamba, 2009). At a local scale in Liwale, the most common features of shifting cultivation observed included use of fire for land preparation, shift of cropping from one field to another and abandonment of fields. Normally, a *shamba* will be cultivated for a couple of farming seasons, then abandoned for a new field as it becomes less productive. In the past, while populations were relatively lower and land availability not an issue, abandoned fields were left to lie fallow for about 25 years (Luoga et al., 2000), allowing enough time for forest regeneration. However with increased population pressure, and thus demand for food, and sometimes growth of produce markets, the fallow period has drastically reduced to less than 3 years (Luoga et al., 2000). The reduced time does not give enough time for forest regeneration, and the higher populations has seen shifts into virgin forests as they seek more productive land. Overall, this may result in forest areas being disturbed. The new clearing practices, sometimes bothering on the use of technology and herbicides as evidenced during the fieldwork, coupled by a shortened fallowing time has resulted in some areas changing from primary vegetation to secondary, and sometimes a total change to grassland (Holden, 2001).

Figure 4.19: Different practices that encourage forest disturbance through degradation and deforestation in Liwale.

The process of a *shamba* preparation, as stated above, varies due to a number of factors and preferences between farmers. Even at local scale, even though the preparation is generally similar, there were variations noted. Even though there are farmers that will deforest an area completely, most farmers cut only the smaller trees into stumps between 50 cm to 1 m high. Thereafter and in successive years, stumps are maintained either by burning them or pruning during field preparation. Grass, herbs and shrubs are completely cleared, and lately some farmers use herbicides, and an example of such a practice is given in Figure 4.20. However, it was observed that the bigger the trunk of the tree, the higher the chances that it was left standing. Depending on the intensity of big trees and a personal preference, some farmers employed methods that would dry the tree, a move usually done to ensure good penetration by sunlight to the crop. The two most used methods for drying the trees is through ring-barking the main trunk of the tree and burning its trunk base (Figure 4.21). A tree may dry the same year if it is ring-barked/burnt, or take a few seasons before it totally dries up, especially for more resistant tree species. For the trees being dried through burning, the process of burning the trunk base is then done successively, every beginning of the ploughing season, during field preparation.

Figure 4.20: A *shamba* where herbicides have been used to cleared herbs in southern Liwale during February 2014.

Figure 4.21: A tree that was dried up through ring-barking and another that had been dried through burning its trunk base.

Local knowledge sourced through discussions with members of the field team (with a rich understanding of the environment and the dynamics of forest disturbance), revealed that the motivation behind cutting trees with trunks smaller in diameter into stumps is not an environmental one. It could have easily been mistaken as being for regeneration purposes. However, farmers do so for ease of weeding. With stumps clearly visible within their crops, it is easy for farmers to plant and weed around them, even predicting where their roots may be, an advantage they lose upon cutting

the whole tree at ground level. When it comes to the large trees, cutting them down burdens farmers, first with chopping and clearing its branches, which may be quite a task depending on the hardness and crown size of the tree. Secondly, as clearing methods are still traditional, mainly being the use of an axe, cutting down and then removing the tree trunk may be either impossible or demanding a lot of manpower, and time consuming. This seems to lead farmers to employing the technique of drying the big trees in their upright position.

Besides directly disturbing the forest for agricultural production, adjacent forests have sometimes suffered from wild fires or fires escaping from areas being cleared for cultivation. Such wild fires are sometimes of very high intensity, as shown by Figure 4.22, which shows fire scars as high up a big tree as the upper stem and branches. In visited sites, trees that had been killed by fires were also evidenced. This underlines a limitation highlighted during the 2013 Southern African Fire Network (SAFNet) meeting that wild fires greatly influenced the frequencies on stand density and species diversity in miombo woodlands like Liwale in Tanzania, yet how fire season and frequency tend to affect vegetation is still insufficiently researched in these woodlands (SAFNet, 2013).

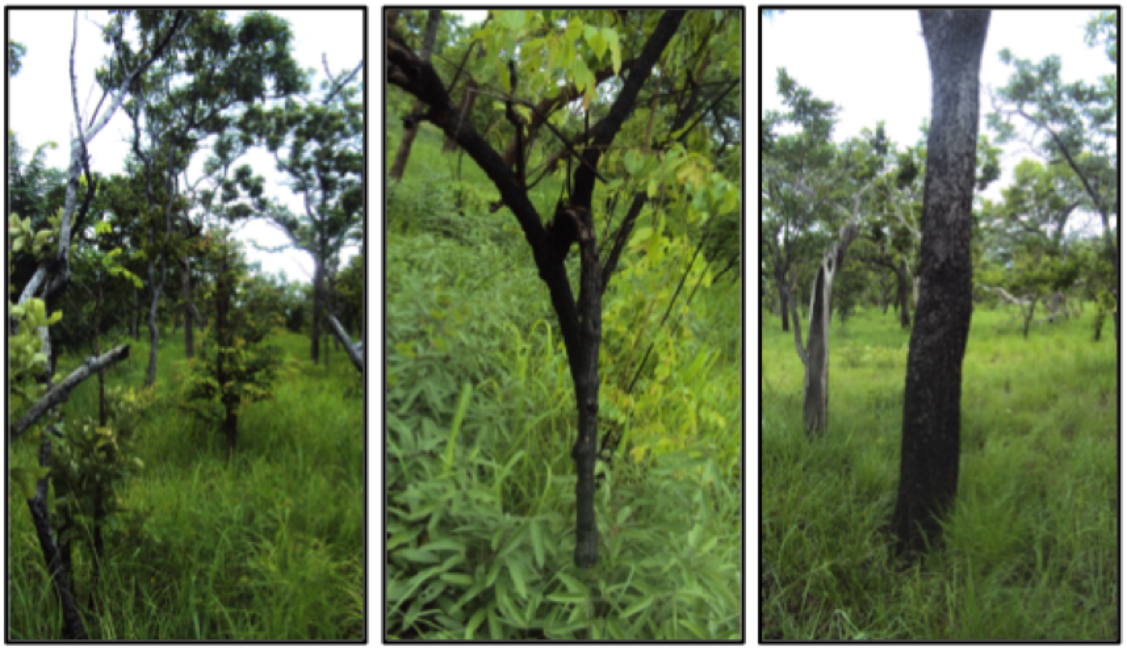


Figure 4.22: An example of trees with evidence of high fire intensity, shown by fire scars observed in the upper part of the tree stem and branches.

4.5 Conclusion

The chapter discussed the different datasets that were used in the study, how and why each was used, and how the data was preprocessed using SPDLib and RSGISLib software. The last section of the chapter detailed on how a fieldwork was undertaken in Liwale, which assisted the researcher to be familiar with the study site, and thus aided the image interpretation process.

Chapter 5

Forest Baseline Classification

5.1 Introduction

Land use has remained a significant driver of habitat degradation, removal and associated biodiversity loss. As a response to forest loss and disturbance, there has been increased global effort to reduce emissions from deforestation and forest degradation, accurate and up-to-date land cover change information has become necessary (Lucas et al., 2014). Land use remains a significant driver of habitat degradation and removal, and associated losses in biodiversity. Such information is key for the understanding and assessment of the environmental effects and consequences of such changes (Shalaby and Tateishi, 2007). A number of remote sensing techniques that have the capability of capturing such changes have been developed. However, the process of extracting the change information from satellite data requires effective and automated change detection techniques, and few reliable techniques have been demonstrated.

Knowledge of the present forest cover distribution and area, especially information on their changing proportions, is crucial to many practitioners and sectors. These include legislators, planners, policy makers, and governmental officials, who if well informed, are better placed to develop good environmental and land use policies. It also assist them to implement effective plans for sustainable development. Overall, it is a complex situation, which highlights the crucial need for accurate, meaningful, current data on LULC and change. Remote sensing-based methodologies for LULC mapping offers robust procedures that are consistent and repeatable. These approaches typically

rely on optical data, but there is considerable value in including radar and LiDAR data (Coppin et al., 2004).

Most parts of the world such as Europe, North America and the tropics across South America, Africa and Asia, have made good advances towards producing forest inventories their forests, and further monitoring them (FAO, 2010c). However, remote sensing approaches to characterising, mapping and monitoring savanna environments are not well advanced, partly because of the complexity of the environment. This is particularly the case in Africa where education and technical capacity is limited.

The system proposed focuses on segmentation of the landscape using a combination of VHR and MR remote sensing data, and the classification of these segments into forest and non-forest classes using a rule-based approach implemented within the EODHaM system proposed by Lucas et al. (2015). Change is then quantified on the basis of reduction or increase in spectral values in relation to the distribution observed for undisturbed and designated classes.

5.2 Background

5.2.1 Classification Systems for LULC Mapping and Monitoring

There are two underlying principles of the performing a classification, and these are *a priori* or *a posteriori* (FAO and UNEP, 2000). *A priori* classification system is an approach based upon defining of classes before the data collection process, and these classes are said to be the abstraction of what is actually occurring. It assumes that all possible classes any user may derive, independent of scale and tools used, are included in the system (FAO and UNEP, 2005). This therefore requires one to deal with all possible combinations of diagnostic criteria before a classification is undertaken. Then during classification, each plot/segment is identified and labelled according to that adopted classification. The advantage with this approach is that one is able to standardize classes independent of the area and the techniques being used, which eliminates bias. However, this means that the approach is very rigid, which leads to problems with some data that may not be easily assignable to any of the pre-defined classes (FAO and UNEP, 2000). In remote sensing terms, this is a supervised approach whereby rule-sets are developed and then used to perform a classification.

On the contrary, *a posteriori* classification uses a direct approach and is exempt from 'preconceived notions' one may have about the area being classified. It only defines classes after data has been clustered using similarities or dissimilarities of a training sample. Its main advantage is its flexibility and adaptability compared to a priori which has been stated to be implicitly rigid. It further ensures that generalization is kept minimal, and thus better fits to collected field observations or training sample of the specific area under study (FAO and UNEP, 2005). On the downside, it tends to be adapted to only the local conditions, which makes it incapable of defining standardized classes. Moreover, the relevance of certain criteria in a certain area may be limited when used elsewhere or in ecologically quite different regions (FAO and UNEP, 2000). Again, in remote sensing terms where EO data are used, this is the unsupervised classification approach.

5.2.2 Shortcomings of Classification Systems

Until the development of the FAOs classification system, there had been no system that was internationally accepted, even though there was a necessity for one. The common cause would be land cover classes that were specific to a particular use or purpose, but inappropriate for other purposes. Another cause relates to the scale used; normally, it would be related to a specific purpose and inappropriate for other purposes. Most classification systems are not perfect, and tend to be area or region-specific. A limitation of a priori systems is that they tend to be generalized whilst posteriori systems are typically localized (FAO and UNEP, 2000). Furthermore, in many cases, the classes used for mapping are arbitrary and particular to the area being studied. However, increasingly there is a need for a systematic and consistent approach to the classification of land covers. Ideally, the classification should be independent of scale. Anderson et al. (1976) observed the complexity associated with LULC and concluded that a perfect classification scheme would never be developed. Reasons for this included the subjectivity of the process, even though an objective numerical approach is often used. In addition, as land cover change was rapid, and was shaped by the demands for natural resources, new categories (e.g., oil palms, sesame plantations) often arose. However, remote sensing, according to (Anderson et al., 1976) provided a platform for the development of a framework that is capable of satisfying the needs of the majority of users.

A common challenge with many classification systems, as noted by FAO and UNEP (2000), is that they are based on natural vegetation, broad land covers or particular features of the landscape (e.g.,

wetlands). Hence, they do not define the full range of possible land covers. For example, vegetation classifications often do not differentiate between agricultural crops whilst agricultural classifications rarely distinguish between different types of natural or semi-natural vegetation.

The range of classifications systems lead to many that are sub-optimal for land cover mapping and monitoring. In particular, there is sometimes bias with some classes given more importance than others. Some classifications may even have imprecise or ambiguous classes and definitions may be absent (Lucas et al., 2015; FAO and UNEP, 2005). In this latter case, there may be insufficient descriptions for the class. Simply describing a class as deciduous forest may give little information on the regeneration or degradation state, the biophysical characteristics (e.g., height, cover or species composition and context within the landscape (e.g., fragmented or contiguous). Legends may therefore only describe a proportion or subset of the entire range of possible classes and the classification scheme may not be compared to those generated in other areas or regions (e.g., in countries neighbouring Tanzania).

Systems should, therefore, be able to distinguish the full range of land cover types and changes between these. In particular, the change classification needs to differentiate between that associated with conversion (from one category to another) or modification (e.g., from pristine natural forests to degraded forests). Many classification systems are too broad, with very few categories describing land covers, and thus do not register all conversions and modifications (FAO and UNEP, 2005).

One system that has the capacity to achieve this is the FAO LCCS classification on which the EODHAM (Lucas et al., 2015) system was developed. The primary advantages of the LCCS classification system can be summarised as follows; a) The biophysical characteristics are inherent to the system. b) The level of fragmentation can be defined. c) Change can be quantified as a change in class or biophysical attribute, including those described by vegetation indices.

The advantage of the FAO LCCS is that criteria relating to biophysical attributes are encompassed, and thus inherent to the system. Other schemes might only consider a component of these. For example, as illustrated in Figure 5.1, a classification based on height and cover (e.g., as obtained using LiDAR and single date optical imagery) would be very different from one based on leaf type (e.g., as obtained using spectral reflectance data, both from single dates and from time-series). The introduction of new criteria, particularly in recent classifications, therefore leads to complexities and differences in the classification process and outputs respectively.

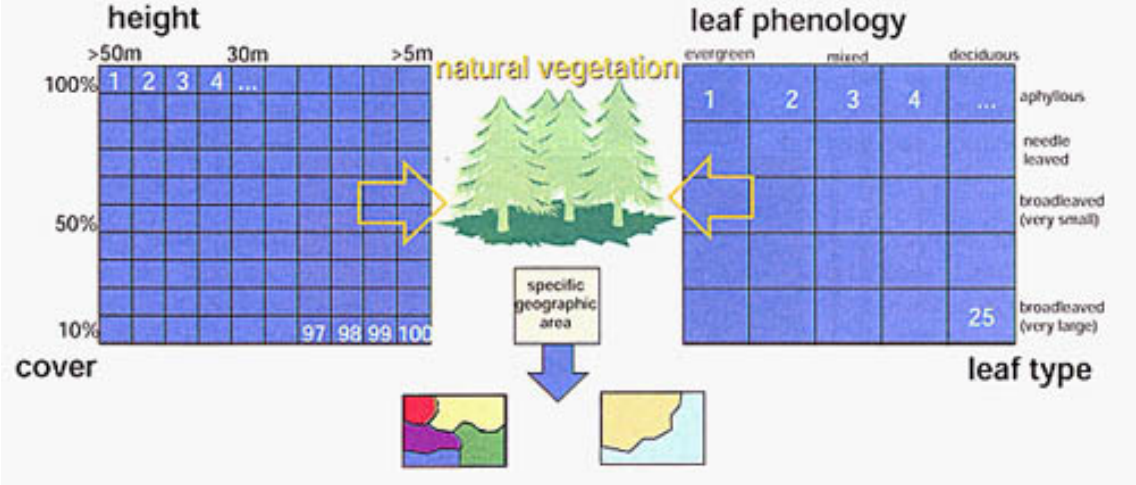


Figure 5.1: Illustration of description of a land cover using two different underlying principles, one using height and cover, and the other using leaf type and phenology.
Source: FAO and UNEP (2005)

5.2.3 Methods for Image-based Data Analysis

The last decade saw an increase in the use of remote sensing for mapping and monitoring forest resources globally. REDD+ aims at improving the MRV for global forest, specifically for the developing world where most forest disturbance is taking place (UNFCCC, 2002; UN-REDD, 2011). In response, the scientific community has done a lot of research aimed at improving image classification procedures, which has resulted in a wide variety of methods for image classification through image interpretation(Warner and Nerry, 2009). Computer-assisted production of land cover maps with adequate spatial details and thematic accuracy from EO satellite image data has continued to challenge the remote sensing research community.

Early studies focusing on land cover classification typically focused on statistical methods and either unsupervised (e.g., K-means) or supervised (e.g., minimum distance or maximum likelihood classifier) algorithms. In each case, spectrally distinct groups of pixels were identified and assigned to a class (Friedl and Brodley, 1997; Zhang et al., 2010b). However, limitations of these techniques related to distributional assumptions and restrictions on a data input (Tseng et al., 2008). Liu and Lathrop (2002) also conveyed that these methods assumed a Gaussian distribution for the training samples for each class but in real life scenario, this was not always the case. Hence, they often struggled to obtain high accuracies, particularly in complex environments. Moreover,

current demand from the remote sensing community and the wider community of end users (e.g. policymakers, foresters, conservationists, ministries, etc.) require ever faster and more accurate classification results. This demand could no longer be fully met using the pixel-based algorithms, mainly due to different characteristics in high resolution imagery and varying user needs (Wang et al., 2009; Blaschke et al., 2014).

The considerable increased availability and advances in high resolution remote sensing data over the recent decades (Rapinel et al., 2014) provided a platform for improved and detailed LULC mapping. However, the first challenge high resolution data gave was that it challenged the use of traditional pixel-based image classifiers. High resolution data has a higher variation in the spectral signatures from land covers compared to lower resolution (e.g., Landsat) imagery. This was recognised by Blaschke (2010); Blaschke et al. (2014) who concluded that this increased variance introduced difficulties in processing data using traditional pixel-based approaches. Hence, greater focus was shifted to the use of object-based methods (commonly referred to as OBIA), as information about objects (e.g. their area, diversity of values, proximity and other relationships to objects) could be explored. These OBIA techniques were even reported to perform above other commonly used classifiers, especially machine learning ones such as ANN and SVM (Zhang et al., 2010b). However, the performance of a classifier is subject to a number of factors, and tend to be environment-specific.

Another issue faced by remote sensing scientists is that of reliably estimating statistical class parameters (Plaza et al., 2009). In many cases, the training samples were too small and the number of features for classification was too high (Foody and Mathur, 2004). This lead to lower accuracies through a process known as the Hughes effect (Hughes, 1968). This stated that a decrease in the training set and an increase in the number of features tends to decrease the classification accuracy. Jimenez and Landgrebe (1998) also demonstrated that higher dimensional spaces are largely empty, whilst Plaza et al. (2009) concluded that low training samples made the estimation processes complicated.

To overcome the limitations of traditional statistical classifiers, research has focused on increasing the accuracy of classification through advancing two main set of techniques (Huang, 2002; Foody and Mathur, 2004). These are evidential reasoning (often referred to as a rule-based or knowledge-based classification) and machine learning (Wijaya and Gloaguen, 2007; Wijaya et al., 2010). All these techniques differ in that they are not based on an assumed parametric model. Hence, the

requirement for a full and representative description of the spectral response of each class was no longer a necessity, and it was sometimes inappropriate in training the classification process (Foody and Mathur, 2004).

5.2.4 Machine Learning Classification Methods

Most mapping monitoring systems are either at regional or national scale, and thus cover large spatial extents, as well as require large volume of remotely sensed data both spatially and temporally. This presents numerous data processing and image interpretation challenges which tend to vary with each environment and the sensor used. Besides the large volume of data to be processed, most complications normally arise from insufficient ground reference data caused by cost and time constraints. The lack of or the limited reference data normally lead to unrepresentative training datasets being used in classifications (Kotsiantis, 2007).

Machine learning algorithms are dominantly non-parametric classifiers that do not assume a specific data distribution to separate a multi-dimensional feature space into classes (Wieland and Pittore, 2014). Machine learning algorithms have been demonstrated as effective for generic LULC mapping (Gopal et al., 1996; Liu and Lathrop, 2002). These algorithms have been preferred due to their ability to handle complex multi-dimensional data with minimal human intervention and reduced processing time compared to the traditional, conventional and parametric classifiers such as maximum likelihood (ML) and nearest neighbour, suggesting their potential suitability in regional-scale forest monitoring (Hansen et al., 2000).

Pixel-based classification studies (Huang, 2002; Pal and Mather, 2003; Dixon and Candade, 2008; Yang et al., 2011; Nitze et al., 2012) showed how non-parametric machine learning algorithms are generally superior over the traditional parametric classifiers (such as nearest neighbour and ML). Using Landsat TM data to perform a LULC classification, Dixon and Candade (2008) found that ANN and SVM achieved similar results, while ML performed significantly worse. The same conclusion was reached by Pal and Mather (2003) when comparing decision tree, ANN and ML algorithms, and reported non-significant differences in the accuracy between the machine learning algorithms, and superiority over the traditional parametric approach. Huang (2002) performed a LULC classification using artificial neural network (ANN), support vector machine (SVM), decision tree (DT) and ML, and concluded that even though the machine learning algorithms outperformed each other in terms of accuracy and the time taken to perform a classification, overall, they did

better than ML.

Decision tree is reported by numerous studies as an efficient tool for image classification and regression problem solving, and these include Civco (1993), Foody et al. (1995), Friedl and Brodley (1997), Huang (2002), Jensen (2005) and Elnaggar and Noller (2010). This method automatically identifies critical variables, and separates them from the rest of not so useful ones for the classification process. Campbell and Wynne (2011) noted that the procedure performs this step without needing much a priori expertise (knowledge). DT also permits the integration of expert knowledge and ancillary data for image analysis and classification, and this includes the process of developing rules through adaptation and learning based on the training data. Pal and Mather (2003) stated that the level of classification accuracy achieved by the classifier was dependent on the training set size, especially for multivariate DT. Moreover, since the tree is constructed from training samples, it tend to suffer from over-fitting upon reaching the full structure.

Mas (2004) is amongst studies that concluded that K-NN performed better than the 'traditional' classification methods. However, another group of studies such as Skidmore et al. (1997) believe that it was not so useful, and especially harder to implement. Another challenge lies with the difficulties that sometimes arise in accounting for how some specific classification outcomes were achieved.

Studies that have used ANN for image classification include Benediktsson et al. (1989), Civco (1993), Schwaiger et al. (1995), Pal and Mather (2003), Tseng et al. (2008), and Plaza et al. (2009). ANNs were reported by Tseng et al. (2008) to produce classifications with similar or higher accuracies from even fewer training samples. It allows for the integration of expert knowledge and ancillary data for image analysis and classification. However, Tseng et al. (2008) further noted that the user must find an appropriate ANNs architecture, which involves a number of parameters. These include a number of hidden nodes, the learning rate, the value of momentum term, the number of training iterations, and specific encoding techniques to represent both input and output information. This usually leads to the adoption of a trial-and-error strategy in order to determine the appropriate parameter values, and thus parametise the model. Secondly, Shin and Lee (2002) highlights the characteristic 'black box' factor of ANNs. With this classifier, a user cannot readily comprehend the final rules that the model acquires. To add, it becomes hard to draw conclusions from the training network with very limited understanding of it. Zhang and Zhu (2011) also observed that ANNs usually converge slowly and tend to converge to a local optimisation.

SVMs are particularly appealing in the remote sensing field due to their ability to generalize well even with limited training samples (a common limitation for remote sensing applications), further often producing higher classification accuracy than the traditional methods (Mantero et al., 2005; Mountrakis et al., 2011). Several studies such as Huang (2002), Mantero et al. (2005), Min et al. (2006), Dixon and Candade (2008), Mountrakis et al. (2011), Nitze et al. (2012) and Moughal (2013) concluded that a SVM can achieve better classification accuracy than other classifiers (both traditional and other machine learning). SVM is particularly effective for the classification of multi-temporal or multi-source data, as demonstrated by studies such as Waske and Benediktsson (2007), Waske and Van Der Linden (2008), Lardeux et al. (2009), Niu and Ban (2010), Hu (2010), and García et al. (2011). Its advantage is in that there is no requirement of the statistical model for the data to be classified. Burges (1998) performed an experiment whereby initially SVM was compared with other algorithms removal of prior knowledge) took place, SVM still out-performed even the best neural networks.

However, other studies such as Otukei and Blaschke (2010) found that DTs performed better than SVMs. Moreover, SVMs suffer from parameter assignment issues that can significantly affect obtained results (Mountrakis et al., 2011). The algorithm requires a training and a testing sample, and the provided training sample determines how good the output classification will be. Alqurashi and Kumar (2013) observed that it is quite sensitive to the amount of the provided training sample, which makes it susceptible to either over-fitting or under-fitting. It also tend to require a long time for training while overall, it is still not as common in remote sensing software (Lu et al., 2004). However, this limitation is slowly being overcome by the move towards open source software.

When reviewing literature with regard to which approach is most commonly used between pixel-based and object-based classification, it reflects that machine learning algorithms have mainly used the former in remote sensing image analysis (Weng, 2012; Wieland and Pittore, 2014). Due to the spatial complexity that is introduced by high resolution remote sensing data, as discussed above, it may be limiting the remote sensing community on what may be achieved using the data. Generally, the application of machine algorithms in the context of an object-based image analysis is lacking (Tzotsos and Argialas, 2006; Wieland and Pittore, 2014). In literature, the machine learning algorithm that has been tested at OBIA level is K-NN, but this too was only compared with the pixel-based ML approach. Most of these found that K-NN outperformed ML, and examples are given below.

Yan et al. (2006) compared pixel-based image analysis using ML and OBIA using K-NN on Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. The study reported that K-NN outperformed ML as an overall accuracy of 83.25% and 46.48% was reported respectively. A similar study by Platt and Rapoza (2008) used multi-spectral IKONOS imagery to compare K-NN and ML, but this time with and without the addition of expert-based knowledge. The results revealed that the K-NN classification best overall accuracy was 78% when using expert knowledge, compared to 64% for the best pixel-based classification using ML (without expert knowledge). Robertson and King (2011) used a two dates Landsat-5 TM images (1995 and 2005) to compare pixel-based and object-based image analysis for classifying broad agricultural land cover types, using ML and k-nearest neighbours (K-NN) algorithms. Their main finding was that the difference in overall accuracy between the approaches was not statistically significant. However, a visual analysis of a post-classification analysis showed that K-NN depicted areas of change more accurately than the ML.

In summary, among their major limitations is their complexity which has still made their operating characteristics to still be poorly understood by both the remote sensing and ecology communities. This continues to limit their potential application in forest mapping and monitoring. Even though their application is becoming more common, there is still a deficiency of knowledge on their capabilities, limitations, and operation for remote sensing applications in ecology (Kavzoglu and Mather, 2003). In general, the comparisons between pixel-based and object-based classifications reveal that object-based outperforms pixel-based when comparing overall classification accuracy using a variety of remotely sensed imagery, across a wide range of settings, from LULC classifications, natural forest mapping, agricultural to urban land cover classification. The biggest challenge with them is their over-reliance on a training dataset, which in this case was a limitation as field data is lacking in Tanzania. Moreover, the difficulty in transferring and applying them at a global scale made machine learning techniques not appealing for this work. Robustness and transferability of the technique advanced for the environment was desired to ensure it can be scaled out nationally and to other environments in the SADC region.

5.2.5 Rule-based Classification Methods

Rule-based (sometimes called knowledge-based) classification is a method that employs a knowledge base to store facts, understandings and heuristics specific to a domain (Warner and Nerry, 2009).

This knowledge base is then used to formulate rules that guide the classification. Most of these algorithms are pixel-based, using spectral dissimilarities, and sometimes ancillary data (Rapinel et al., 2014). A critical part of the rule-based methods is the development of rule-sets to describe the classes of interest. Often threshold values need to be manually adjusted and classification rules are strongly bound to a particular image scene and image type. Only a few studies have designed generalized rule-sets that aim at being transferable (Drgu et al., 2014).

The commonly adopted approach when using rule-based classification, especially for high resolution EO data is to perform it at object based image analysis (OBIA) level than at pixel-based level. Compared to the pixel-based approach, the object-oriented approach classifies objects in images from spectral criteria, and can also incorporate texture, context and shape criteria (Drgu et al., 2014). Two studies (Burnett and Blaschke, 2003; Bock et al., 2005) pointed out that the object-oriented approach allows for the analysis of images in a multi-scale framework that is suited to hierarchical typology for vegetation mapping. In addition, Lucas et al. (2015) states that the approach permits for the integration of thematic layers, which provides contextual information. However, according to (Blaschke, 2010) and Rapinel et al. (2014), there is still a need for the development of OBIA algorithms as there are still few of them, and these are not easy to implement nor are they transferable compared to per-pixel classification techniques.

Developed rule-based OBIA approaches have been reported to perform better and achieve higher accuracies than pixel-based classifications. For rule-based classifications, the OBIA algorithms still mainly rely on the use of the spectral properties of the image to produce rules which are then used together with other criteria such as shape, texture and context to perform the classification.

At OBIA level, as stated by Blaschke (2010) and Rapinel et al. (2014) that more work still need to be done, studies that have compared machine algorithms to rule-based classification approaches are lacking. As shown in Table 5.1, the last few years has seen a lot of object-based classification studies. Studies comparing machine learning and rule-based techniques would give an insight on how these fare against each other. However, as the review shows, these rule-based studies were comparing pixel-based to object-based techniques. A rule-based and machine learning algorithm (decision tree) that was done by Tseng et al. (2008).

Myint et al. (2011) used Quickbird imagery to classify urban land cover. They compared results from a ML pixel-based classification with an object-based classifier using K-NN and a series of fuzzy membership functions. The object-based classification (90.4%) outperformed the pixel-based

classification (67.6%) in overall accuracy for the original image, but the difference between the techniques in a a test image was reduced to less than 10% (95.2% and 87.8%, respectively). K-NN assigns the image objects to the classes of interest according to their similarity to selected training samples and defined feature space (indices, texture, spectral information). However, its dependence on the training samples makes it less transferable to other images (Belgiu and Drgu, 2014).

Fuzzy rule-based classification was used by Bardossy and Samaniego (2002), Lucas et al. (2007), Lucas et al. (2011), and Zhang and Zhu (2011) at pixel-based level, while Sebari and He (2013) used it at object-based level. This algorithm also uses a rule system derived from a training set (Bardossy and Samaniego, 2002),through the use of an automatic knowledge extraction technique to gather knowledge from the training samples (Kartikeyan et al., 1994). According to Table 5.1, for a pixel-based classification for land cover comparing ML and a fuzzy rule-based approach, ML outperformed the rule-based approach with an overall accuracy of 98.7% and 94.4% respectively (Bardossy and Samaniego, 2002). At object-based level, even though using very high resolution IKONOS images, the best overall accuracy the study could achieve was 80% (Sebari and He, 2013).

For both the fuzzy logic and K-NN, the reliance on the training sample introduces the same limitation suffered by machine learning algorithms, with their dependence on a training samples making them less transferable to other images and environments (Belgiu and Drgu, 2014). In the case of Liwale, and Africa generally where field data may be very limited, if any, the site was without a training sample. Moreover, the goal was to develop techniques that could be scaled out to other sensors, and especially to other similar environments in the region. These limitations of the K-NN and fuzzy rule-based approaches meant that these two techniques would not be best suited for the research.

Most studies in literature still used the traditional rule-based approach, and as shown in Table 5.1, some achieved high overall accuracies. Studies such as Lucas et al. (2007), Kahya et al. (2010), Lucas et al. (2011), Zhang and Zhu (2011), and Yang et al. (2015), used the rule-based pixel-based algorithm, and reported results with overall accuracies satisfying the requirements for practical applications. In Beijing, China, Zhang and Zhu (2011) achieved an overall accuracy of 90.91% and a kappa coefficient of 0.8341 for land cover classification. The results showed that forest land, standard water bodies, farmland and buildings can be extracted effectively (over 90% accuracy),

whereas non-standard water bodies, roads and shadows could only be extracted slightly less effectively (over 80%). This was attributed to the high percentage of forest cover and the interlaced distribution of roads, shadows and forestland. These types of land use are relatively fragmentary, an attribute of the savanna woodlands.

Huang and Jensen (1997) used SPOT HRV data in combination with other ancillary data to show that the rule-based approach was superior over maximum likelihood and unsupervised classification in classifying for land cover. Lucas et al. (2011) developed rules through effectively capturing the knowledge and expertise of field ecologists and remote sensing scientists in eCognition Definiens software to implement an object-orientated rule-based classification of detailed habitats at a national level in Wales. This produced the first national map of habitats through object-oriented classification from EO data in Europe, with an overall accuracy of <80%, and between 70% and 90% for most classes.

Suchenwirth et al. (2012) found that the overall accuracy achieved for vegetation types was higher for a combination of spectral- and knowledge-based classification than when spectral values were used in only, the inclusion of additional knowledge improved the results. The study further concluded that generally, the more knowledge rules applied, the better the classification accuracy, with reduced potential errors. A new approach that still need more research, especially in different environments, is the use of a combination of the machine learning and rule-based algorithms, as was used by Niu and Ban (2010). The study found that applying a rule-based post-classification to an SVM classification improved the classification, yielding an overall accuracy of 90% and kappa coefficient of 0.89.

A recent study by Rapinel et al. (2014) trialled an integration of both the pixel-based and object-based for natural vegetation mapping, using high resolution WorldView-2 data. The study found that the classification achieved a kappa coefficient of 0.9. A similar approach was used by Lucas et al. (2015) using WorldView-2 data, which through the use of a combination of pixel and object-based procedures. The primary source of the rule sets used on most of the studies, both pixel-based and object-based was based on knowledge base which enabled effective use of the spectral data of the image. These were then supplemented with other ancillary data, which in most of these studies was reported to enhance the overall accuracy. What can be drown from Table 5.1 is that for those studies on high resolution imagery that compared a pixel-based and an object-based algorithm

The motivation behind the use of a object-oriented rule-based approach based only on the spectral properties of the imagery is that the Liwale site is relatively rural, with minimal other land use or cover types other than those related to vegetation. Studies such as Lewinski and Bochenek (2008) and Zhang and Zhu (2011), showed how the integration of other knowledge such as textural properties and shape enhanced the results. However, most such studies were studying urban and semi-urban environments, and even then, concluded that the land cover classes relating to forests and farm/agricultural land were best separated using their characteristic vegetation spectral signature. The African savanna environments are too irregular, heterogeneous and fragmented, making properties such as object shape not appropriate.

In general, satellite image distributions could differ in combinations of partially uncontrollable or unknown factors, including atmospheric effects, sensor gain or seasonal variations in the image scene. A question of transferability of classifiers on an extended and partly domain-invariant feature space derived from an object-based image analysis remains still largely unanswered in remote sensing applications (Wieland and Pittore, 2014), and this work sought to further explore this area by developing the classification technique on a RapidEye data, and further trial it on another sensor image type (Landsat). Landsat has been used by studies such as Bardossy and Samaniego (2002), Lucas et al. (2007), Tseng et al. (2008), and Kahya et al. (2010), but all these used pixel-based classification techniques. However, RapidEye does not seem to have been used that often for LULC classification, especially using the rule-based approach. With other REDD nations such as Guyana having adopted a national monitoring system that uses RapidEye data, it is worth exploring if the same results being achieved in other areas can be achieved in the savanna forests of Africa.

Table 5.1: Literature on studies that have used the rule-based approach for a range of applications in remote sensing.

Application	**Sensor(s)**	**Pixel- or object-based**	**Overall accuracy**	**Kappa coefficient**	**Study**
Natural vegetation mapping	WorldView-2	Hybrid (both)		0.9	Rapinel et al. (2014)
Urban vegetation mapping	Quickbird	Pixel- vs object-based	67.6% (Pixel-based) 90.4% (Object-based)		Myint et al. (2011)
Urban land cover mapping	IKONOS	Object-based	80%		Sebari and He (2013)
Urban land cover mapping	WorldView-2	Object-based	71% - 87% (For four sites)		Belgiu and Drgu (2014)
Land cover classification	Landsat	Pixel-based	98.7% (ML) 94.4% (Fuzzy)		Bardossy and Samaniego (2002)
Land cover classification	Landsat	Pixel-based	95%		Kahya et al. (2010)
Land cover mapping	SPOT	Object-based	89.1%	0.87	Lewinski and Bochenek (2008)
Land cover classification	IKONOS	Object-based	60% (Spectral) 70% (Knowledge-based)	0.53 0.64	Suchenwirth et al. (2012)
LULC classification	Landsat	Pixel-based	84.9%		Lucas et al. (2007)
LULC classification	WorldView-2	Pixel- & object-based			Lucas et al. (2014)
Land cover classification	Landsat	Pixel-based	73.3% (Rule-based) 72.1% (Decision tree)		Tseng et al. (2008)
LULC classification	CBERS	Pixel-based	82.8% (Single date image) 87.9% (Multi-temporal images)		Yang et al. (2015)
LULC classification	Quickbird	Object-based	90.91%	0.83	Zhang and Zhu (2011)

5.2.6 FAO Land Cover Classification System (LCCS)

The LCCS was been developed by the FAO and the UNEP as a comprehensive, standardized a priori classification system, to meet the need for improved access to reliable and standardized information on LULC change (FAO and UNEP, 2005). LCCS created to meet specific user requirements, and created for mapping exercises, independent of the scale or means used to map. It was designed to overcome, or at least minimize the limitations of other classification systems as discussed above. It enables a comparison of land cover classes regardless of data source, thematic discipline or country, making it a globally applicable classification system. The LCCS enhances the standardization process and minimizes the problem of dealing with a very large amount of pre-defined classes. To ensure standardization amongst users, and to facilitate the complex classification process and ensure standardization, users are guided to choose the appropriate class down the hierarchy during classification.

The greatest difficulty in the mapping of the LCCS classes is the differentiation of aquatic vegetation as well as cultivated and managed areas (Lucas et al., 2015). However, the advantages of the classifier/approach are manifold, as argued by FAO and UNEP (2005).

Advantages include that its legend is applicable to any land cover and can be consistently applied. Furthermore, it provides the platform for quantifying change. FAO and UNEP (2005) states that the approach design ensures that only the most appropriate classifiers are used, while inaccurate classifier combinations are avoided. It is a programme that guides the user step-by-step (classifier by classifier), assisting them to choose the correct class. Even though maintaining standardization, it is still as flexible, enabling a user to create classes at different levels and to use optional modifiers, environmental attributes and specific technical attributes in combination.

The LCCS has an initial Dichotomous Phase, wherein eight major land covers are defined. This is achieved by first defining the extent of vegetated and non-vegetated land covers as well as those that are terrestrial or aquatic (Level 1). These are then cross tabulated to form a new set of classes (Level 2), such as terrestrial vegetated or aquatic non-vegetated. Subsequently, the landscape is categorised according to whether it is natural/semi-natural or artificial/cultivated, with this then cross tabulated with Level 2 categories. The resulting classes at Level 3 include, for example, natural/semi-natural terrestrial vegetated, artificial bare (e.g., urban) or natural water (FAO and UNEP, 2005).

The second phase is termed the Modular Hierarchical Phase. Here, the components of land cover types (e.g., vegetation height, cover, phenology and leaf type) are defined and the codes assigned to each (in Level 4) are combined to generate a classification. Through this approach, numerous elements of the landscape are defined which gives considerable detail to the final classification (FAO and UNEP, 2005).

The system has already been used to generate the global forest resource assessments undertaken by the FAO in 2000, 2005 and 2010. It contributes towards ensuring that data are harmonised and standardised, both nationally and globally. This way, it enables the comparison of land cover classes regardless of data source, economic sector or country (FAO, 2010c). Accuracy assessment of the end product can be generated by class or by the individual classifiers forming the class.

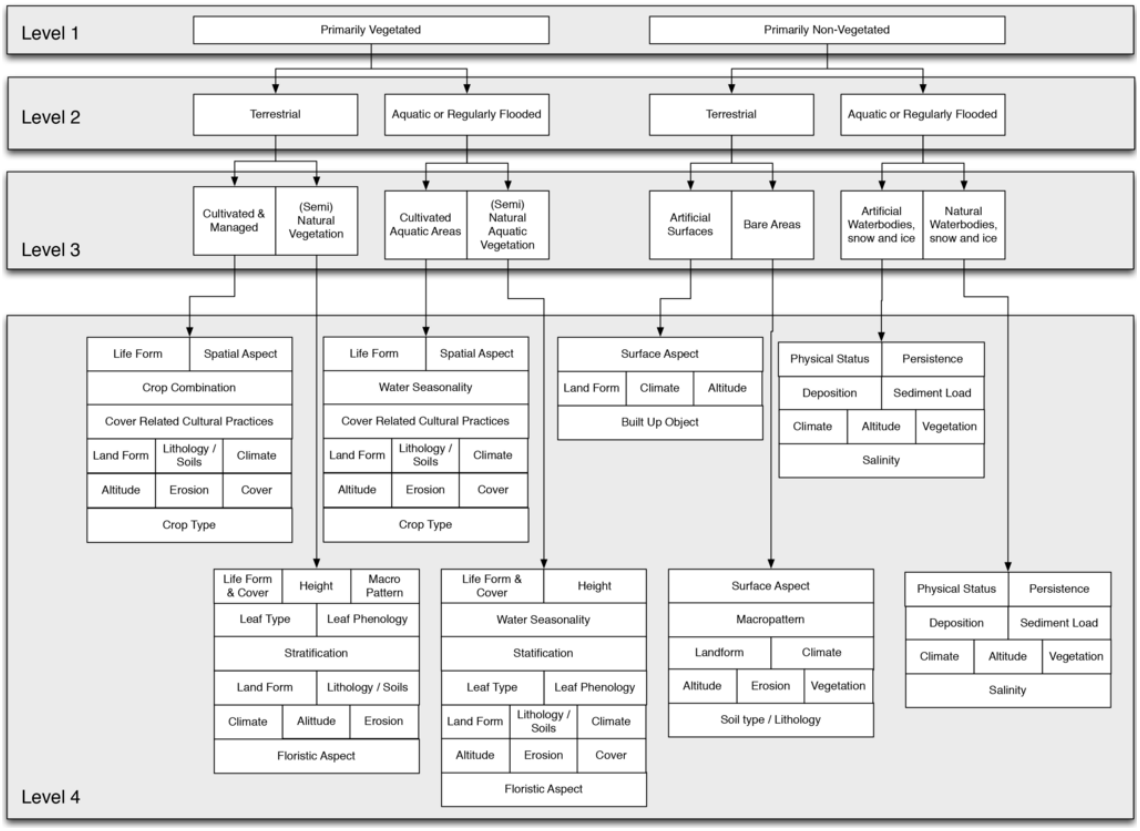


Figure 5.2: The hierarchical structure of Levels 1 to 4 of the Land Cover Classification Scheme (LCCS).

Source: FAO and UNEP (2005)

5.2.7 EODHaM Classification System

Appreciating the ability and advantages of the LCCS classification system and its robustness, the BIOdiversity multi-SOurce monitoring System: from Space TO Species (BIO_SOS) project (EU FP7 Funded) used the FAOs LCCS taxonomy to develop the Earth Observation Data for HAbitat Monitoring (EODHaM) system (Lucas et al., 2015). This is a semi-automated classification system aimed at improving consistent mapping and monitoring of biodiversity. A component that is considered very key is the inclusion of decision rules within a hierarchical classification structure. These rules are generated from the expert knowledge of ecologists and RS scientists. It translates mapped classes to General Habitat Categories (GHCs), and uses a combination of pixel and object-based procedures. EODHaM is performed in three major stages. The first and second stages are focused on the use of earth observation data alone with expert knowledge to generate classes according to the LCCS taxonomy (Levels 1 to 3 and beyond). The third stage then translates the final LCCS classes into GHCs from which Annex I habitat type maps (EU Habitats Directive) are derived (Lucas et al., 2015).

Having been adapted from the LCCS classification, and thus possessing a very robust nature that makes it transferable to other environments globally, this research modified the EODHaM system detailed in Lucas et al. (2015), and applied it for LULC classification of the savanna woodlands of Liwale, Tanzania. Adopted from the universally applicable LCCS, the system was used to perform a classification from Level 1 to Level 3 of the LCCS classification, with the last step being the grouping of classes into either forest and non-forest in order to establish the forest baseline.

5.3 Methods

The adopted approach from modifying the EODHaM system (Lucas et al., 2015) for the image classification process involved a number of steps, summarised in Figure 5.3. The main steps included first calculating vegetation indices, and stacking them with the RapidEye image. Then the new image was segmented, and the resulting RAT image populated with the statistics of the objects. Thereafter, with the assistance of field points mapped during the fieldwork, unique rulesets were developed for each class at the different levels of classification. This was then followed by running the LCCS classification, classifying for Levels 1 to 3, and finally grouping the land classes at Level 3 into either forest or non-forest.

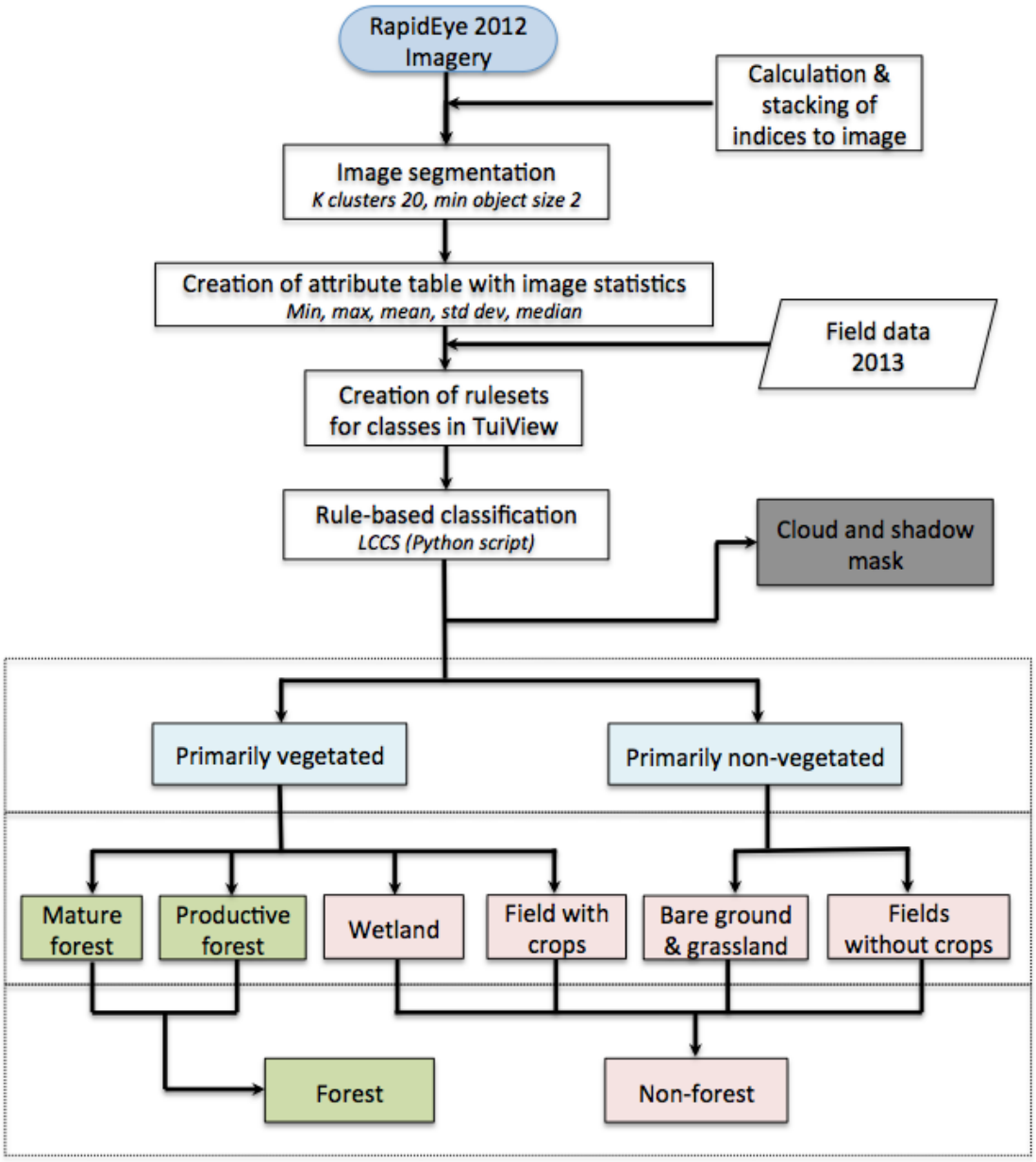


Figure 5.3: Schematic overview of the developed approach for classifying land covers adopted from the rule-based LCCS classification for RapidEye data.

These steps are explained in detail in the next subsections. Worth noting is that during the implementation of the classification, it was found that classes were too few, with barely any areas in the site that were aquatic in nature, except for wetlands. Therefore, Level 2 was intentionally skipped, and the classification moving straight from Level 1 to Level 3 as shown in Figure 5.3. This is a commonly used approach in Literature, and Belgiu et al. (2014) is an example of a study

that used the image objects generated at Level 1 to classify for land cover classes.

5.3.1 Image Mosaic and Vegetation Indices

For the RapidEye data, as outlined in Chapter 4, 11 relatively cloud-free (<20%) scenes were acquired on the 12^{th} May 2012. From these, an image mosaic was generated and from it, several spectral indices were derived (Table 5.2). These indices were associated with the biophysical attributes of land surfaces, including productivity and water amounts (Jensen, 2000). The range of values and the normalisation value for each is given in Table 5.3. In preparation for landscape segmentation, the normalised indices were then stacked into one image with the original bands of the RapidEye image mosaic to produce a new image.

Table 5.2: Indices trialled out during classification, and the respective formula for each of these indices.

Index	**Index abbr**	**Formula**
Blue Green Pigment Index	BGI	Blue - Red
Blue Red Pigment Index	BRI	(Blue - Red)/(Blue + Red)
Canopy Chlorophyll Content Index	CCCI	(NIR - RedEdge)/(NIR + RedEdge)
Chlorophyll Green	ChlGreen	(NIR/Green) - 1
Chlorophyll Index RedEdge	CIRedEdge	(NIR/RedEdge) - 1
Chlorophyll Vegetation Index	CVI	NIR*(Red/(Green*Green))
Forest Discriminant Index	FDI	NIR - (RedEdge + Blue)
Leaf Chlorophyll Index	LCI	(NIR - RedEdge)/(NIR + Red)
Normalized Difference Vegetation Index	NDVI	(NIR - Red)/(NIR + Red)
Plant Senescence Reflectance Index	PSRI	(Red - Blue)/RedEdge
Red-Edge Position Linear Interpolation	REP	RedEdge - (NIR - Red)
Water Band Index	WBI	(Blue/NIR)
Woody Vegetation Ratio	Woody	(Blue - Green)/(Blue + Green)

Table 5.3: The range of each computed index, and the normalisation value used to move the data range away from negative and zero values.

Index	**Index abbr**	**Min value**	**Max Value**	**Normalisation value**
Blue Green Pigment Index	BGI	-231	67	500
Blue Red Pigment Index	BRI	-0.9	0.339	100
Canopy Chlorophyll Content Index	CCCI	-0.244	0.526	100
Chlorophyll Index Green	ChlGreen	-382	434	500
Chlorophyll Index RedEdge	CIRedEdge	-0.393	2.223	100
Chlorophyll Vegetation Index	CVI	30.911	841.807	0
Forest Discriminant Index	FDI	-763	271	900
Leaf Chlorophyll Index	LCI	-0.237	0.637	100
Normalized Difference Vegetation Index	NDVI	-0.286	0.863	100
Plant Senescence Reflectance Index	PSRI	-0.929	0.971	100
Red-Edge Position Linear Interpolation	REP	-255	1150	500
Water Band Index	WBI	0.011	1.758	100
Woody Vegetation Index	Woody	-0.904	0.141	100

5.3.2 Image Segmentation

In the classification of high resolution remote sensing data, object-based image analysis (OBIA), which involves first segmenting the image, is reported to give better accuracies than pixel-based approaches (Bouziani et al., 2010). Image segmentation is the process of splitting up the image into homogeneous and spatially contiguous image objects (Suchenwirth et al., 2012). The segmentation was performed on the new image stack, using an algorithm of Bunting et al. (2014a), which was implemented in RSGISLib software. The algorithm is repeatable in that it always yields the same result given the same input image and parameters (Clewley et al., 2014), and its steps described in Figure 5.4. It uses an implementation of K-means clustering to generate seeds. Once pixels have been assigned to the associated cluster centre, the clumps are iteratively eliminated if they are below the minimum mapping unit threshold to the neighbouring clump that is closest in colour, defined through Euclidean distance. After the elimination process, the final clumps are relabelled for consecutive numbering, and no data regions are assigned a value of zero (Bunting et al., 2014a).

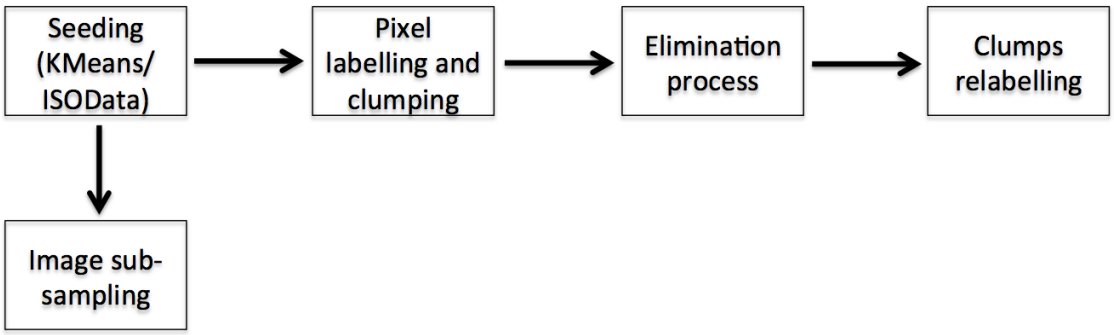


Figure 5.4: A flowchart of the segmentation algorithm implemented in RSGISLib.

The algorithm employs two key parameters; number of clusters (k) in the K-means which are used to seed the K-means algorithm, and the minimum object size in pixels below which objects are eliminated. Other parameters include distance threshold (which prevents merging), sampling of the input image for input into the K-means, and the maximum number of iterations within K-means. The size of segments is controlled by the k value.

According to Pradhan et al. (2010), the K-means clustering is one of the simplest unsupervised learning algorithms. The procedure follows a simple and easy way to classify a given data set through a certain fixed number of clusters (k clusters). For k clusters, k centroids are defined,

one for each cluster. Since the positioning of each of these centroids affects the result, one has to find the optimum positioning for these, and Pradhan et al. (2010) advices that these centroids are best placed as far from one another as possible. Thereafter, each pixel or object is associated with the nearest centroid, followed by re-calculating the new positions of the centroids of the clusters resulting from the previous step. These two steps of assigning pixels/objects to the nearest centroid and recalculating the position of the resultant centroids are iterated until the centroid position remains similar (within certain thresholds of movement from the cluster mean). As a result, objects are separated and grouped based on data properties, namely; spectral reflectance and derived indices.

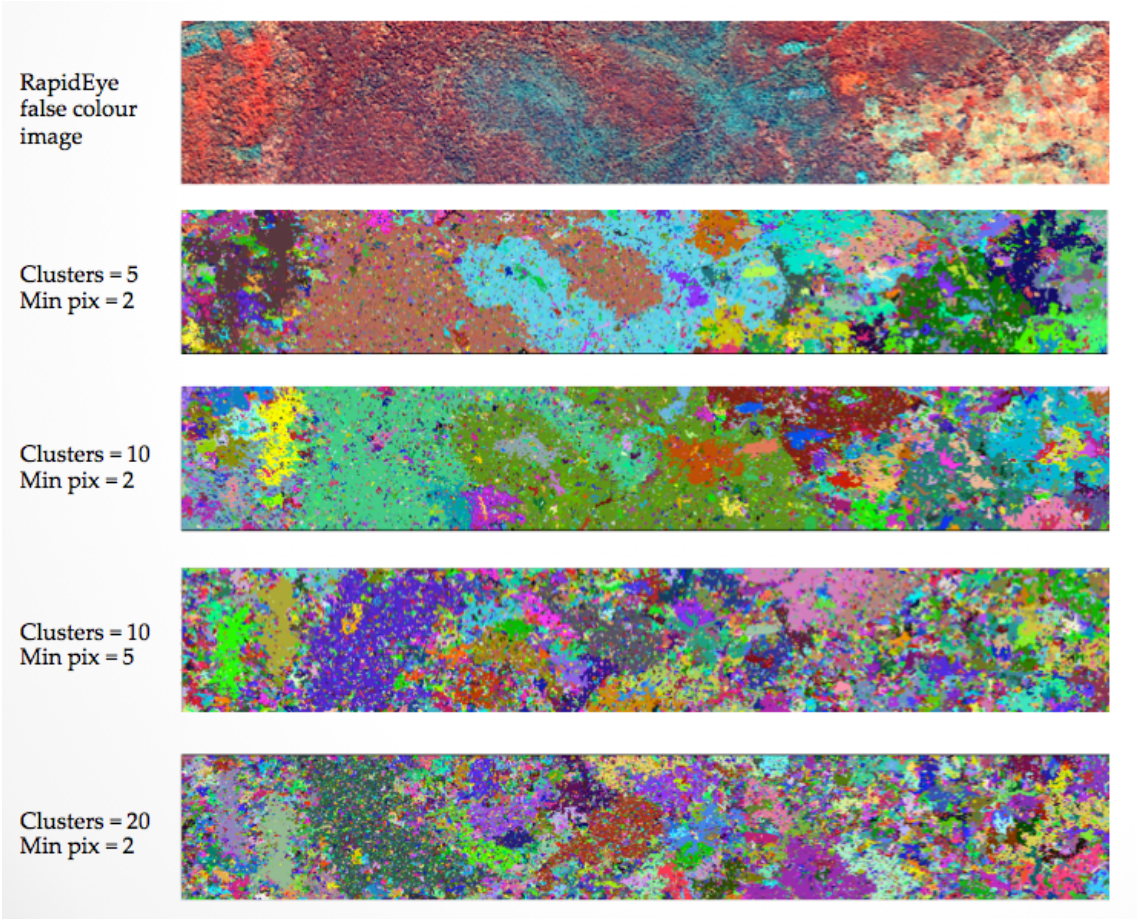


Figure 5.5: Examples of different segmentation results from using different parameters within the segmentation algorithm.

For the RapidEye data, the optimal segmentation was achieved using a k value of 20 and a minimum object size of 2 pixels. Figure 5.5 shows a segmentation result from the optimum parameters and

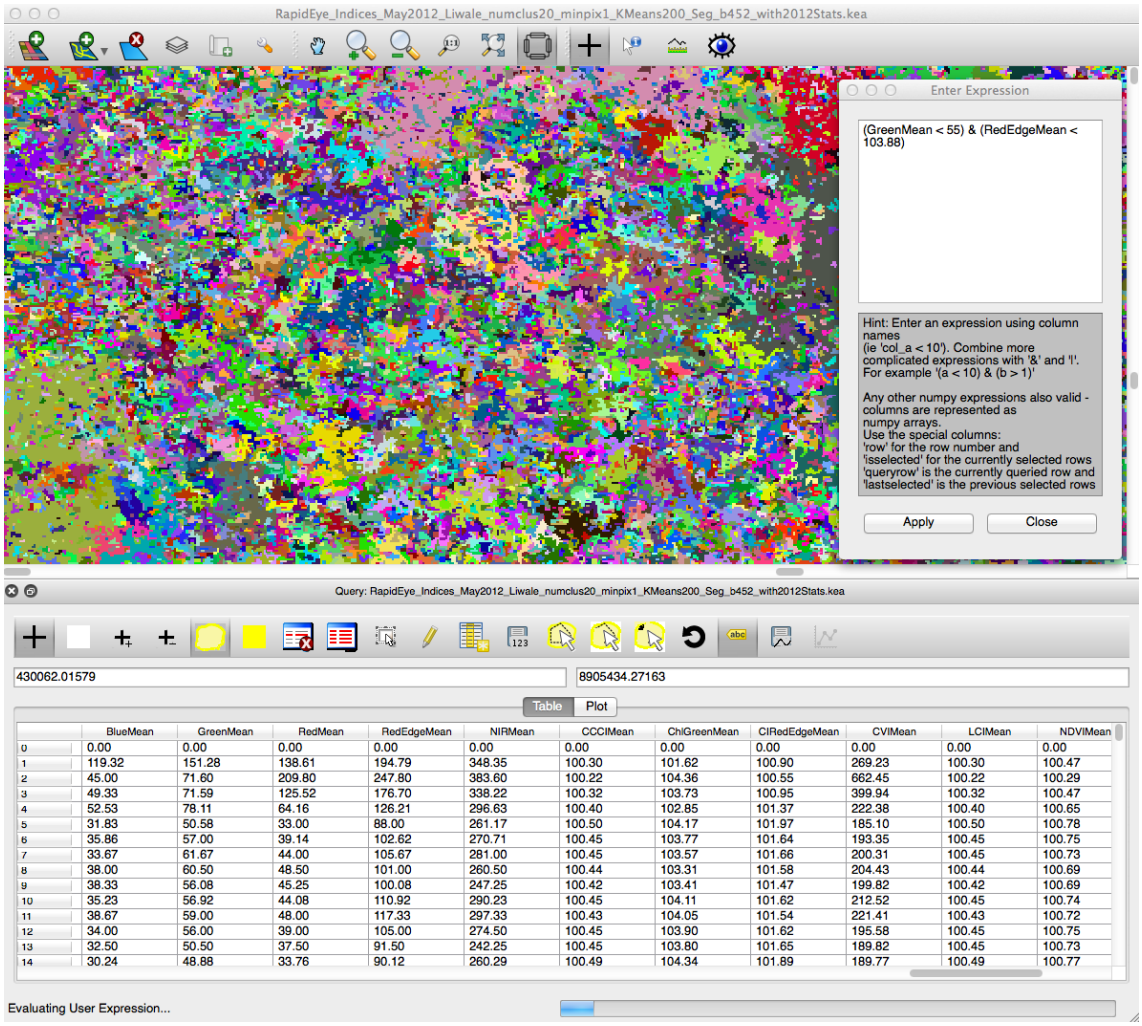
other parameter combinations.

5.3.3 Attribute Table Creation

The resultant product from the segmentation process was a raster attribute table (RAT) image made of objects. The RAT was then populated with the spectral and contextual attributes of these objects before the classification to LCCS categories. Each clump/object was attributed with the image statistics (minimum, maximum, sum, mean, standard deviation, median and count), area, and position (Clewley et al., 2014). All these were stored within the RAT image as an attribute table.

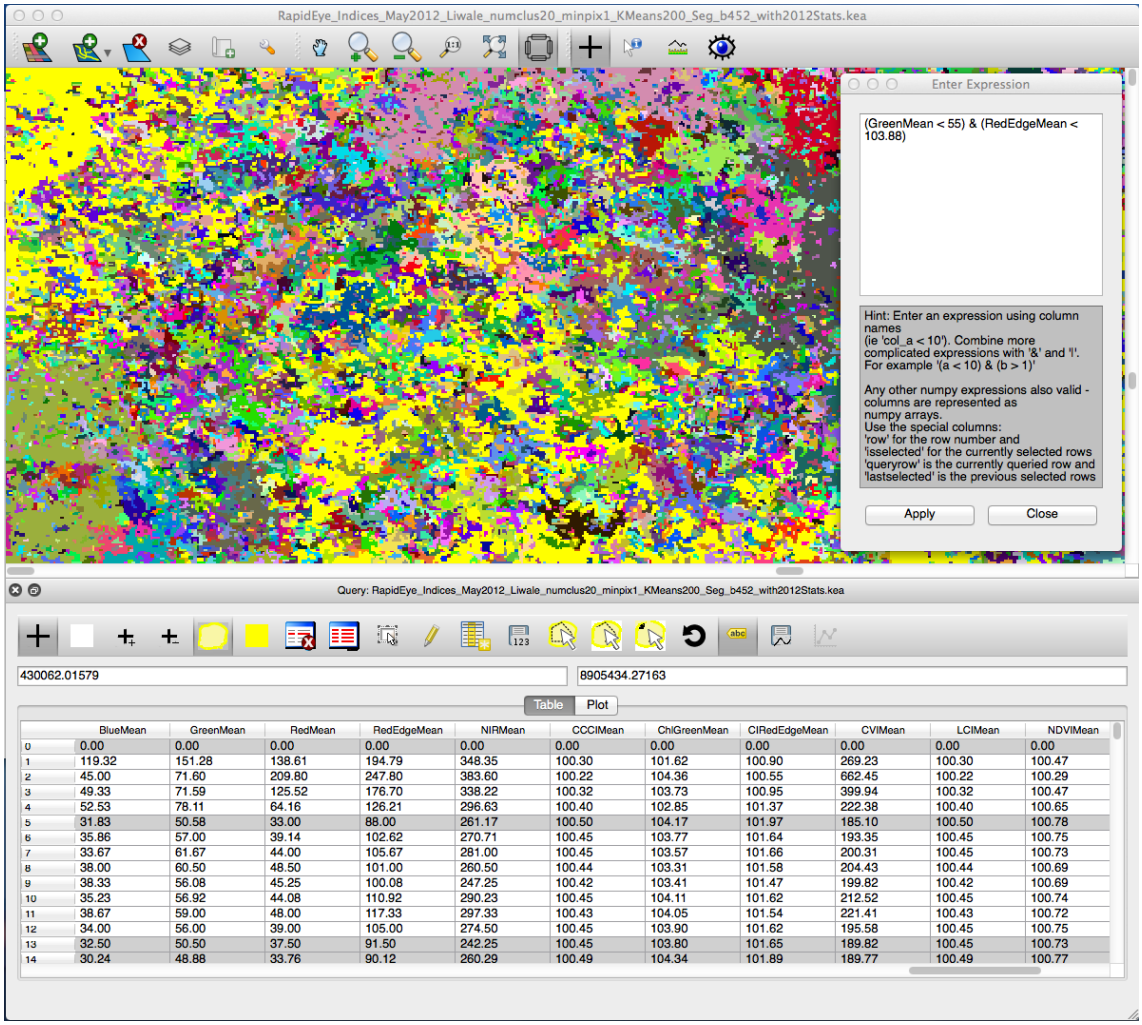
5.3.4 Creation of Rulesets for Classes

Using the statistics that had been populated into the attribute table, a rule-based classification was undertaken. A Raster Input/Output Simplification (RIOS) library was used to read columns and write new columns within a Raster Attribute Table (RAT). The procedures were undertaken in Python and columns were represented as NumPy arrays, an extension to Python for handling large multi-dimensional arrays (Bunting et al., 2014a). The advantage of NumPy is that it supports a number of functions (e.g. mathematical operations). This permitted the development of unique rules for each of the land classes encountered. This was assisted by using the TuiView Interface using the 'Select using Expression' tool (Figure 5.6 and Figure 5.7). The NumPy 'where' statement was used for 'if-else' statement.



	BlueMean	GreenMean	RedMean	RedEdgeMean	NIRMean	CCCIMean	ChlGreenMean	CIRedEdgeMean	CVIMean	LCIMean	NDVIMean
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	119.32	151.28	138.61	194.79	348.35	100.30	101.62	100.90	269.23	100.30	100.47
2	45.00	71.60	209.80	247.80	383.60	100.22	104.36	100.55	662.45	100.22	100.29
3	49.33	71.59	125.52	176.70	338.22	100.32	103.73	100.95	399.94	100.32	100.47
4	52.53	78.11	64.16	126.21	296.63	100.40	102.85	101.37	222.38	100.40	100.65
5	31.83	50.58	33.00	88.00	261.17	100.50	104.17	101.97	185.10	100.50	100.78
6	35.86	57.00	39.14	102.62	270.71	100.45	103.77	101.64	193.35	100.45	100.75
7	33.67	61.67	44.00	105.67	281.00	100.45	103.57	101.66	200.31	100.45	100.73
8	38.00	60.50	48.50	101.00	260.50	100.44	103.31	101.58	204.43	100.44	100.69
9	38.33	56.08	45.25	100.08	247.25	100.42	103.41	101.47	199.82	100.42	100.69
10	35.23	56.92	44.08	110.92	290.23	100.45	104.11	101.62	212.52	100.45	100.74
11	38.67	59.00	48.00	117.33	297.33	100.43	104.05	101.54	221.41	100.43	100.72
12	34.00	56.00	39.00	105.00	274.50	100.45	103.90	101.62	195.58	100.45	100.75
13	32.50	50.50	37.50	91.50	242.25	100.45	103.80	101.65	189.82	100.45	100.73
14	30.24	48.88	33.76	90.12	260.29	100.49	104.34	101.89	189.77	100.49	100.77

Figure 5.6: A screenshot of the 'Select using expression' tool in TuiView as an expression was being executed.



	BlueMean	GreenMean	RedMean	RedEdgeMean	NIRMean	CCCIMean	ChlGreenMean	CIRedEdgeMean	CVIMean	LCIMean	NDVIMean
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	119.32	151.28	138.61	194.79	348.35	100.30	101.62	100.90	269.23	100.30	100.47
2	45.00	71.60	209.80	247.80	383.60	100.22	104.36	100.55	662.45	100.22	100.29
3	49.33	71.59	125.52	176.70	338.22	100.32	103.73	100.95	399.94	100.32	100.47
4	52.53	78.11	64.16	126.21	296.63	100.40	102.85	101.37	222.38	100.40	100.65
5	31.83	50.58	33.00	88.00	261.17	100.50	104.17	101.97	185.10	100.50	100.78
6	35.86	57.00	39.14	102.62	270.71	100.45	103.77	101.64	193.35	100.45	100.75
7	33.67	61.67	44.00	105.67	281.00	100.45	103.57	101.66	200.31	100.45	100.73
8	38.00	60.50	48.50	101.00	260.50	100.44	103.31	101.58	204.43	100.44	100.69
9	38.33	56.08	45.25	100.08	247.25	100.42	103.41	101.47	199.82	100.42	100.69
10	35.23	56.92	44.08	110.92	290.23	100.45	104.11	101.62	212.52	100.45	100.74
11	38.67	59.00	48.00	117.33	297.33	100.43	104.05	101.54	221.41	100.43	100.72
12	34.00	56.00	39.00	105.00	274.50	100.45	103.90	101.62	195.58	100.45	100.75
13	32.50	50.50	37.50	91.50	242.25	100.45	103.80	101.65	189.82	100.45	100.73
14	30.24	48.88	33.76	90.12	260.29	100.49	104.34	101.89	189.77	100.49	100.77

Figure 5.7: A screenshot of the 'Select using expression' tool in TuiView, with features meeting the expression selected in yellow colour, and appearing grayed out on the attribute table.

The advantage with this approach is that it allows the user to visually assess and easily adapt rules being developed (Clewley et al., 2014). To do so, the NumPy syntax is used within TuiView (Clewley et al., 2014), which gives the user the chance to develop rules using GUI interface.

5.3.5 Complexity of Liwale Environment and the Advantage of the Rule-based LCCS Approach

The taxonomy used for land cover classification was the FAO's LCCS, with land cover subsequently classified to either forest or non-forest classes. Clewley et al. (2014) noted that this classification

could be implemented using any remote sensing data acquired at any scale, and that generated classes were consistent and globally relevant.

As discussed above, different classification approaches differ in the steps taken (Pradhan et al., 2010). The three basic classification approaches are broadly grouped into supervised, unsupervised and hybrid classification techniques. Supervised classification techniques, as used in the EODHaM system (Lucas et al., 2015), requires training areas or rules to be defined by the user, and these are then used to determine properties of each class being classified for. On the contrary, an unsupervised classification approach searches for natural groups of pixels (clusters) within the dataset by means of assessing the relative locations of the pixels in the feature space. Thereafter, the user interprets the clusters and assigns them to the respective land classes. More techniques are being developed, and these combine the advantages of both the supervised classification and unsupervised classification to produce a better hybrid classification approach (Tademir et al., 2012; Rapinel et al., 2014).

For Liwale, it was found that these woodlands are very complex and heterogeneous in nature. Figure 5.8 illustrates numerous land classes within a small feature space, with some of these classes appearing similar. To ensure minimal confusion between classes and to keep the classification errors minimal in the forest baseline that was to be used for change detection, the LCCS-based EODHaM classification system (Lucas et al., 2015) was the preferred option. Using expert knowledge of the environment, and having benefited from a fieldwork to the site, the sought classes were well known and understood. This was a great benefit in the development of rulesets used to discriminate the land classes of interest, and these would be adjusted until the desired objects representing a class of interest were well captured. Using a learning machine approach does not avail a user this advantage, and as discussed above, these algorithms can produce very complex results that are hard for the user to interpret. And for such a complex environment as shown in Figure 5.8, such a limitation was likely to be experienced.

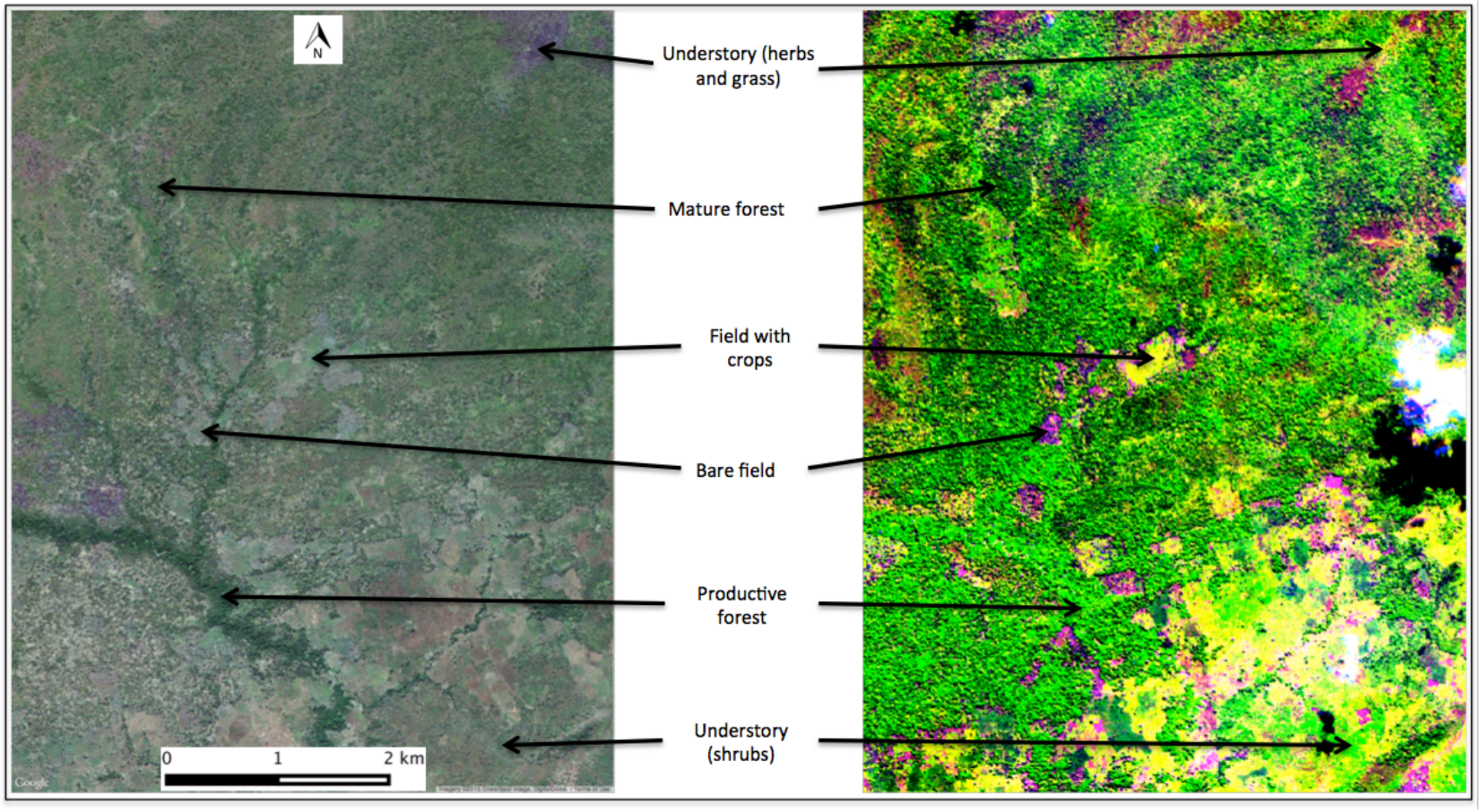


Figure 5.8: A false colour image of Liwale (RedEdge, NIR, Blue) showing different LULC classes that rulesets were developed for, and classified for in Liwale.

In addition, for these techniques to yield good results, the user has to provide a very good training set that covers the whole range of the class. Under this environment, this was always going to be hard to achieve.For example, as illustrated in Figure 5.9, three distinct kinds of the landscape were fields for the local subsistence farmers. One kind was without crop, then two both had crop on them, yet having a different spectral properties when displayed as a false colour image (B:RedEdge, G:NIR, R:Blue). This could be attributed to different crops having been cultivated, and to the crops being at different stages of growth. One of the two was confusing with mature forest as in the false colour image, it was displayed in dark red, a reflectance associated with mature forest. Such complexity was best dealt with and controlled for using rule-sets than a machine learning approach at this stage.

Furthermore, since the idea was to develop a robust method that can be transferred to other savanna forests and other environments in this deciduous African environment, it was a better option to base it on rule-sets that a user can easily adjust for their environment than a machine learning algorithm that may result in over-fitting in a different spatial location, or even the same area if an image differed due to seasonal effects, or for a different EO sensor.

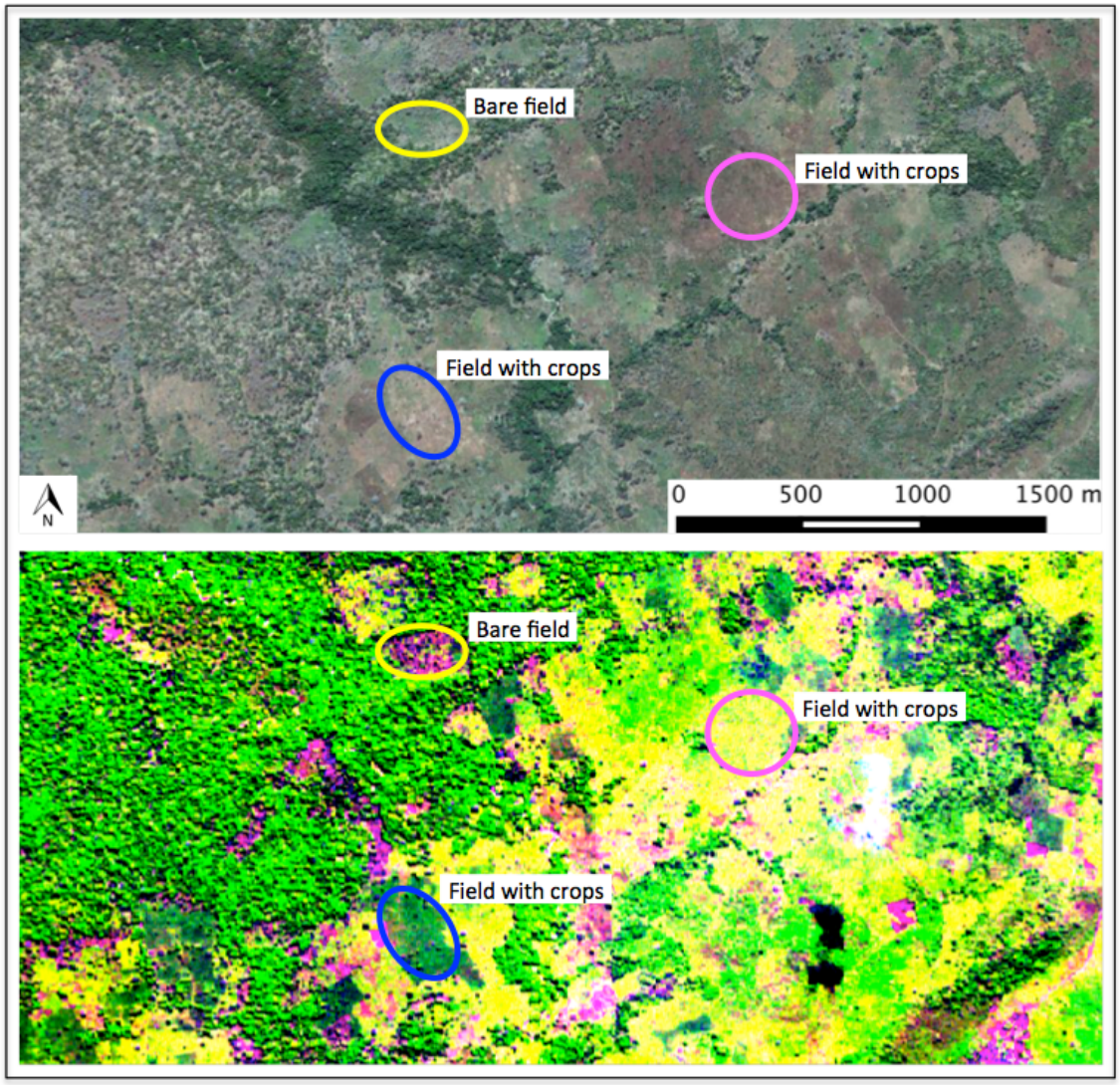


Figure 5.9: A false colour image of Liwale (B:RedEdge, G:NIR, R:Blue) showing different crop fields in the same environment in Liwale. This illustrates the likelihood of other approaches such as machine learning producing very complex classifications which would be difficult to interpret.

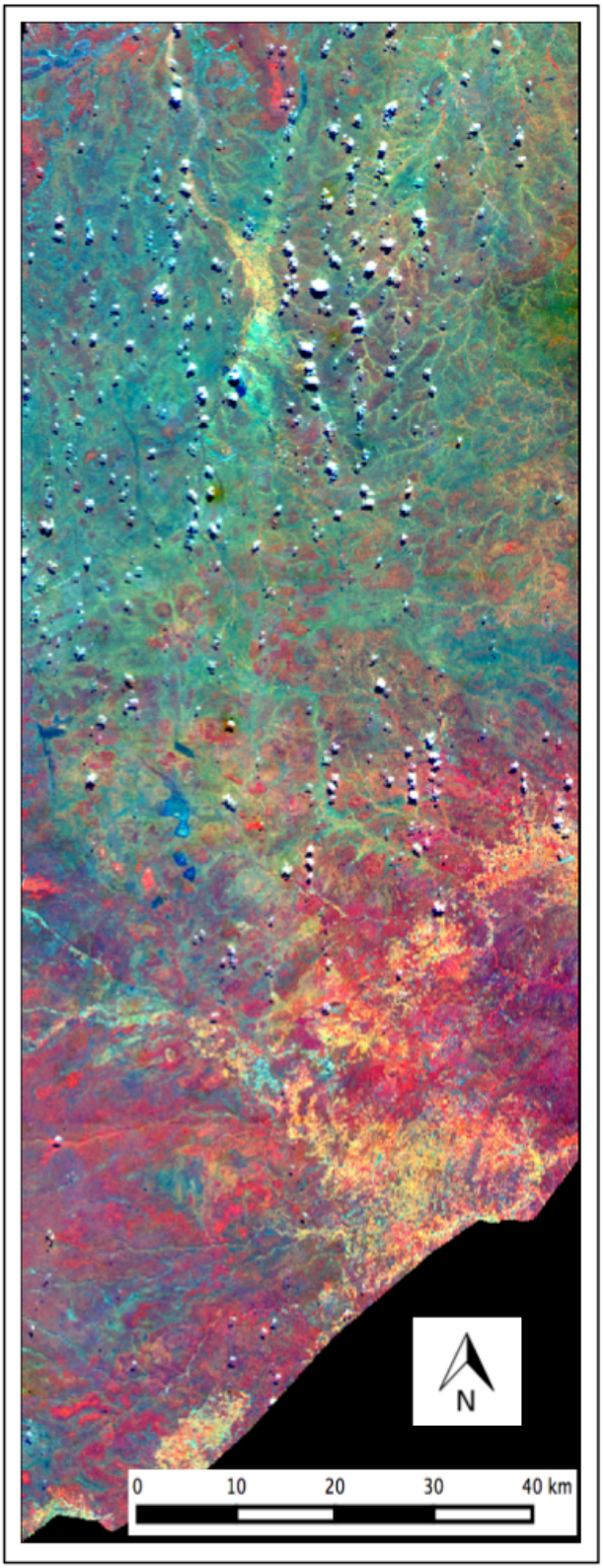


Figure 5.10: False colour RapidEye image (B:NIR, G:RedEdge, R:Blue) of Liwale for May 2012, showing the spectral difference of the forested area in the northern and southern parts of the site.

From Figure 5.10, the spectral difference of the forest class across the site can be seen. The lower southern part the site is dominated by closed forest while the northern part is mainly open woodlands. The northern part had a forest environment that is more fragmented, which may also account for the difference in the spectral properties, and thus the establishment of a unique rule-set for the mature forest for each. However, difference was only noted for the RedEdge spectral band, with the threshold values for the other bands (blue and green) being applicable for both the northern and southern environments.

5.3.6 The Rulesets and the Classification

The first set of rulesets that were established were those that were used to classify the objects made of pixels covered by clouds and shadows. These were then used to mask out the 'no data' areas. The study opted for the manual technique for cloud and shadow masking as it was found to be sufficiently quick and effective for the needs of the study. Importantly, it lined up with the use of rulesets, which was being proposed for the baseline mapping process. In addition, there are similar studies that used this manual approach in literature, an example being Jin et al. (2013), that used thresholds derived from spectral information from the blue, short-wave infrared, and thermal infrared bands of Landsat imagery to detect and mask out clouds and shadow from the imagery. To note, however, is that there are potential ways to automate this process, and examples include otsu thresholding, morphological feature extraction, Bayesian theory and FMASK algorithm (Zhu and Woodcock, 2012; **?**; Xu et al., 2014; Karlsson et al., 2015; Zhu et al., 2015). These more complex algorithms would be more appropriate for an operational system than the manual approach.

This step was followed by the development of a set of rules for classifying for primarily vegetated and primarily non-vegetated classes at Level 1 of the LCCS. Lastly, a set of rules was developed to classify for land classes (Level 3) using the image objects generated at Level 1. Table 5.4 is a summary of the unique rule-sets that were derived for each of the respective classes at the Modular Hierarchical Phase of the classification. Beyond this phase (at Level 3), land classes were grouped into either forest or non-forest.

To note is that the relatively forested areas and settled areas in the south displayed different reflectance values compared to the less forested areas (savannas) in the north. Hence, the site was divided into a northern part and a southern part, and a sightly different rule (105 versus 103.88 respectively for the mean RedEdge reflectance) was applied for the mature forest class. Table 5.4

also shows that the field with crops class was hard to develop a rule-set, and required the use of numerous spectral bands. It was mainly confused with the productive forest class. Another class that was found to be confusing with the productive forest class during the development of the rule-set was some grassland that were also rich in herbaceous plants.

Table 5.4: The developed rule-sets for the different classes and levels of the LCCS that were used to establish the forest baseline of Liwale from RapidEye data.

Level	**Class**	**Ruleset**
No data mask	Cloud	Red Max >150 NIR Min >150 NIR Max >280
	Shadow	Green Mean >61 & <77 NIR Mean <200
LCCS Level 1	Primarily non-vegetated	NDVI Mean <100.6
	Primarily vegetated	NDVI Mean >100.6
LCCS Level 3	Mature forest (Southern Liwale)	Green Mean <70 RedEdge Mean <103.88
	Mature forest (Northern Liwale)	Green Mean <70 RedEdge Mean <105
	Productive forest	CIRedEdge Mean <102.30 Red Mean >100
	Wetland	CIRedEdge Mean <101.15 Red Mean >40
	Field with crops	LCI Mean <100.51 RedEdge Mean >55 & <122 Green Mean >55 NIR Mean <330
	Field without crops	RedEdge Mean <122 Green Mean <55
	Bare ground and grassland	RedEdge Mean >100 Blue Mean >30 Red Mean >50

Following the establishment of the rule-sets for each class of interest at the respective levels (Table 5.4), a python script for the classification was then written. Thereafter, the classification script for application, a forest/non-forest classification was undertaken.

5.4 Results

5.4.1 Forest Baseline from LCCS Classification

The site of Liwale that had appropriate data, as discussed under the 'RapidEye Used Data' section, had a total land area of 6529 km^2, of which 819.42 km^2 did not have data due to cloud cover and cloud shadow, as shown in Figure 5.11. These areas were therefore masked out. Out of 5709.58 km^2 remaining area, 2088.94 km^2 was classified as forest, while 3620.64 km^2 was non-forest (Table 5.5).

Table 5.5: The forest baseline of Liwale for the year 2012.

Class	**Area (km^2)**
No data	819.42
Non-forest	3620.64
Forest	2088.94
Total	6529.00

The forest class was made up of mature and productive forest classes. The non-forest class consisted of all the other kinds of LULC classes, such as fields, grasslands, shrubs, and bare ground. Since the aim was to establish the forest baseline of the area, these were grouped into either forest or non-forest as shown in Figure 5.3.

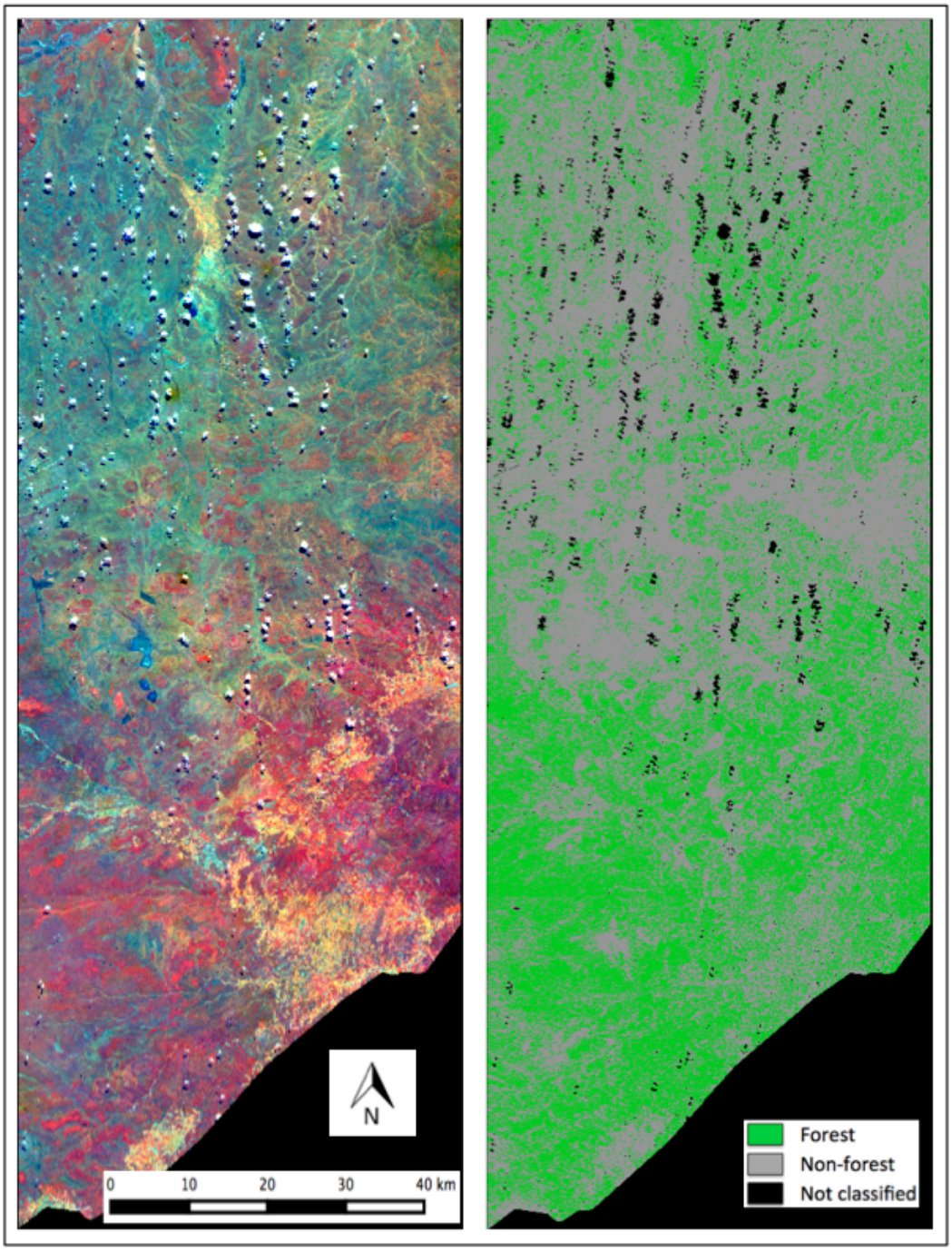


Figure 5.11: A false colour image (B:NIR, G:RedEdge, R:Blue) and the forest/non-forest classification of the whole site of Liwale.

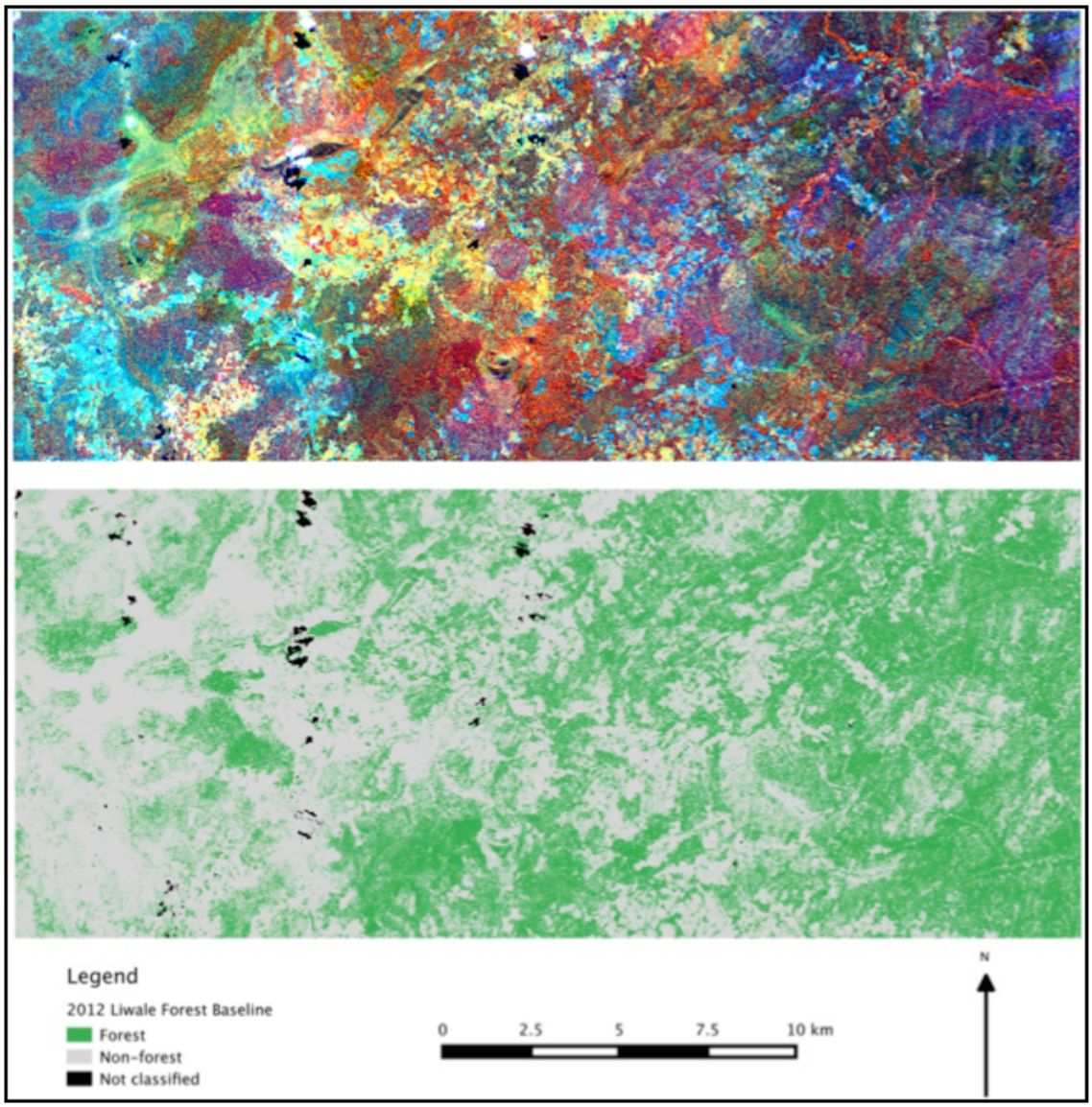


Figure 5.12: A false colour image (B:NIR, G:RedEdge, R:Blue) and the forest/non-forest classification of Liwale SLU site.

Figure 5.12 and Figure 5.13 are a subset of the SLU site. This part of the Liwale site is an area that was specifically chosen because there is a lot of human disturbance taking place within the eastern part of it, while the western section of it is relatively undisturbed. Comparing the classification achieved for the site, against the false colour image of Liwale, the light blueish non-vegetated (non-forest areas) have been well classified.

Overall, the LCCS-adapted EODHaM classification system yielded a good forest baseline for the

site, even though due to numerous environmental and data limitations, there were some classes that had error. Figure 5.14 is an example of different land classes that were well classified. Polygon B (blue) shows a mix of forest and shrubs, polygon A (yellow) shows a block of predominantly forested patch of land which however does not look so obvious to the eye as it looks to be more bare ground than vegetated. This has been well classified for, and the trees (woody) have been picked up very well on the classification. Polygon C (purple) shows a section that is dominated by agricultural fields. Some of these fields have productive crops and shrubs that are relatively similar to productive forest. As it can be seen on the map, these classes were well separated. This part of the site is made up of relatively closed forest, and there was less trees shadow effect, which could be attributed to the better resultant segmentation, and ultimately classification results.

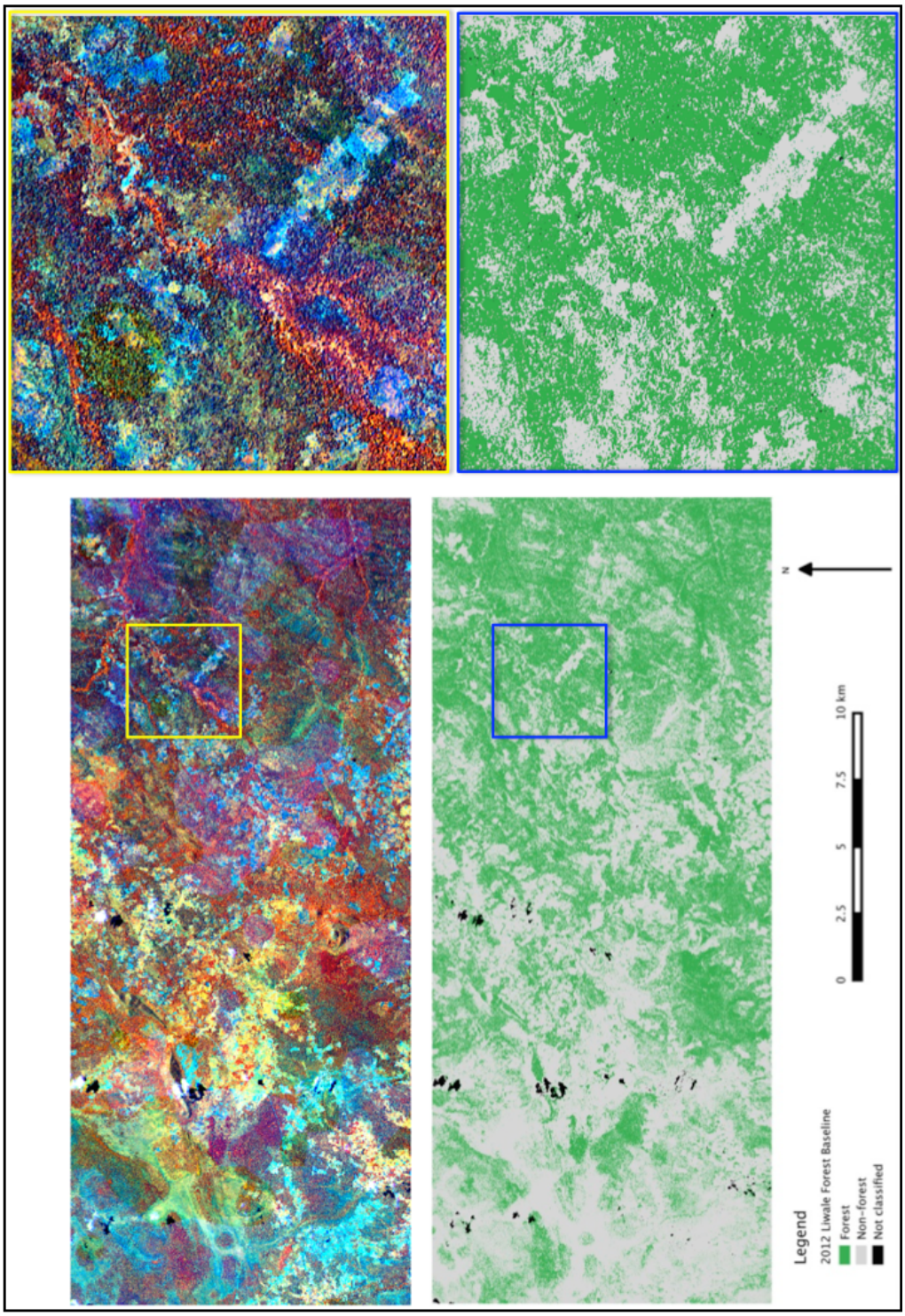


Figure 5.13: A subset of the produced baseline for Liwale SLU, and a false colour image (B:NIR, G:RedEdge, R:Blue).

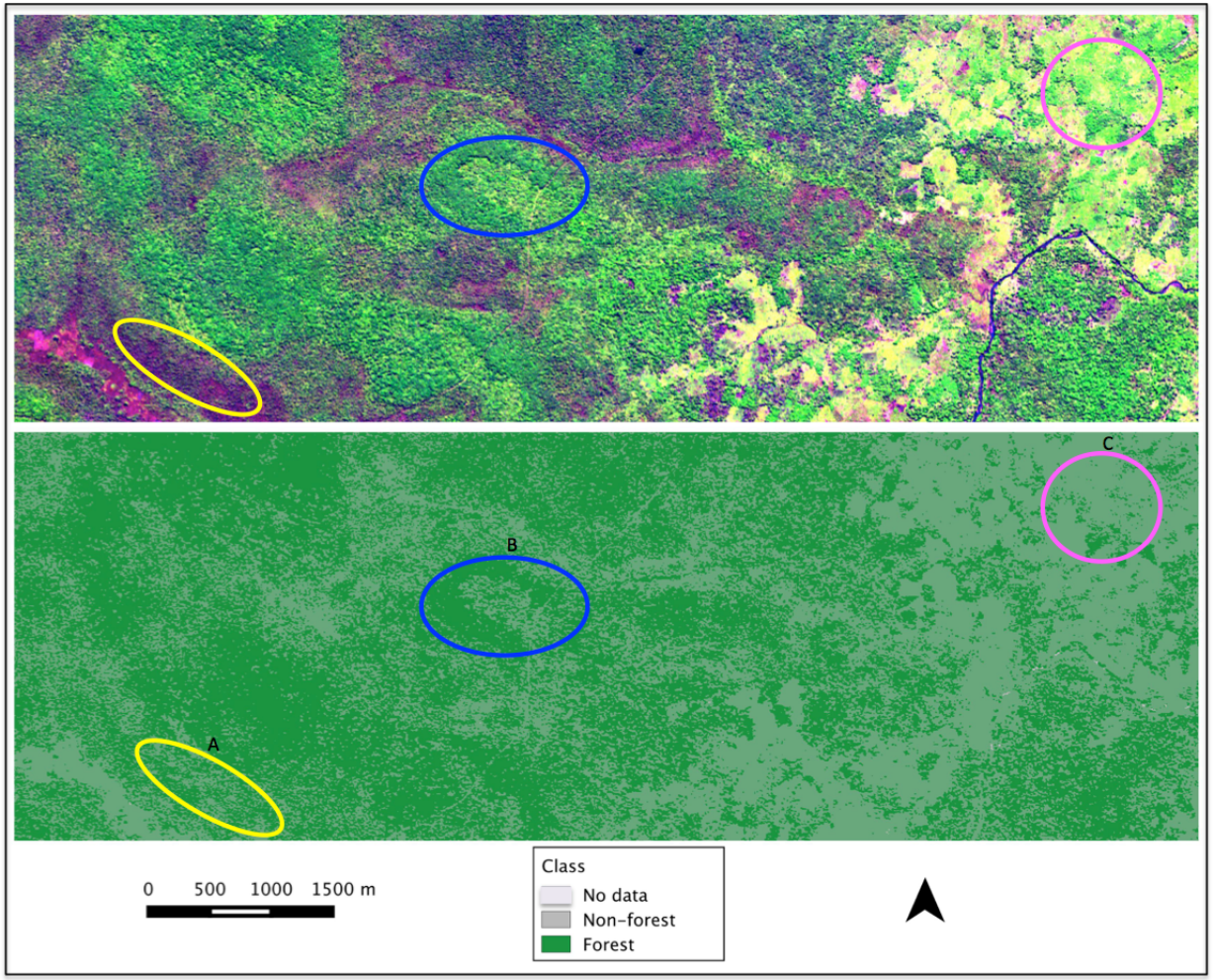


Figure 5.14: Example of LULC classes that were well classified, showing that the LCCS technique is capable of producing a good forest baseline. RapidEye false colour image was displayed as B:RedEdge, G:NIR, R:Blue.

But as shown in Figure 5.15, some parts of the site were affected by shadows. This part of the site is more sparsely vegetated and there is a lot of shadow effect on the image, which led to some segments around shadows being poorly demarcated. Individual trees within a grassland or agricultural fields were affected most, as it is illustrated in Figure 5.15 by the blue polygons (D and E). However, these trees seemed hardest to pick in bare fields than in fields with crops or grassland areas. In this illustration, polygon D is located in a mosaic shrubby vegetation and fields with crops, which have three trees spaced out. These were well classified. However, due to shadow effect, their boundaries may not be fully correct. On the contrary, polygon E shows a number of trees spaced out in a bare field, and these have been missed.

It was also hard to correctly demarcate, and thus classify the edges different classes, especially forest edges, mainly due to the effect of shadow. The yellow polygons (A, B and C) show situations where even though it can be seen that it is a forest block boundary or that there is a cluster of trees, one cannot make the actual edges of the forest block due to the heavy shadow on these boundaries. As it can be seen within these circles, eventually the shadow which indicates that there is a tree is what was classified, which was believed not to be the full tree cover. The red polygon (F) shows an area that has a number of classes sharing boundaries, and these are productive forest, shrubs and fields with crops. Even though the fields with crops were well separated and classified as non-forest, the same did not hold true for the understorey vegetation (shrubs) as some parts of it was classified as forest. If shrubs are thick and have good canopy cover, these can have a spectral signature very similar to productive forest. Since optical data does not have height information, it becomes very hard to separate these two if it is being used in isolation, as was the case in this study. It is one of the limitations optical data are known to suffer from.

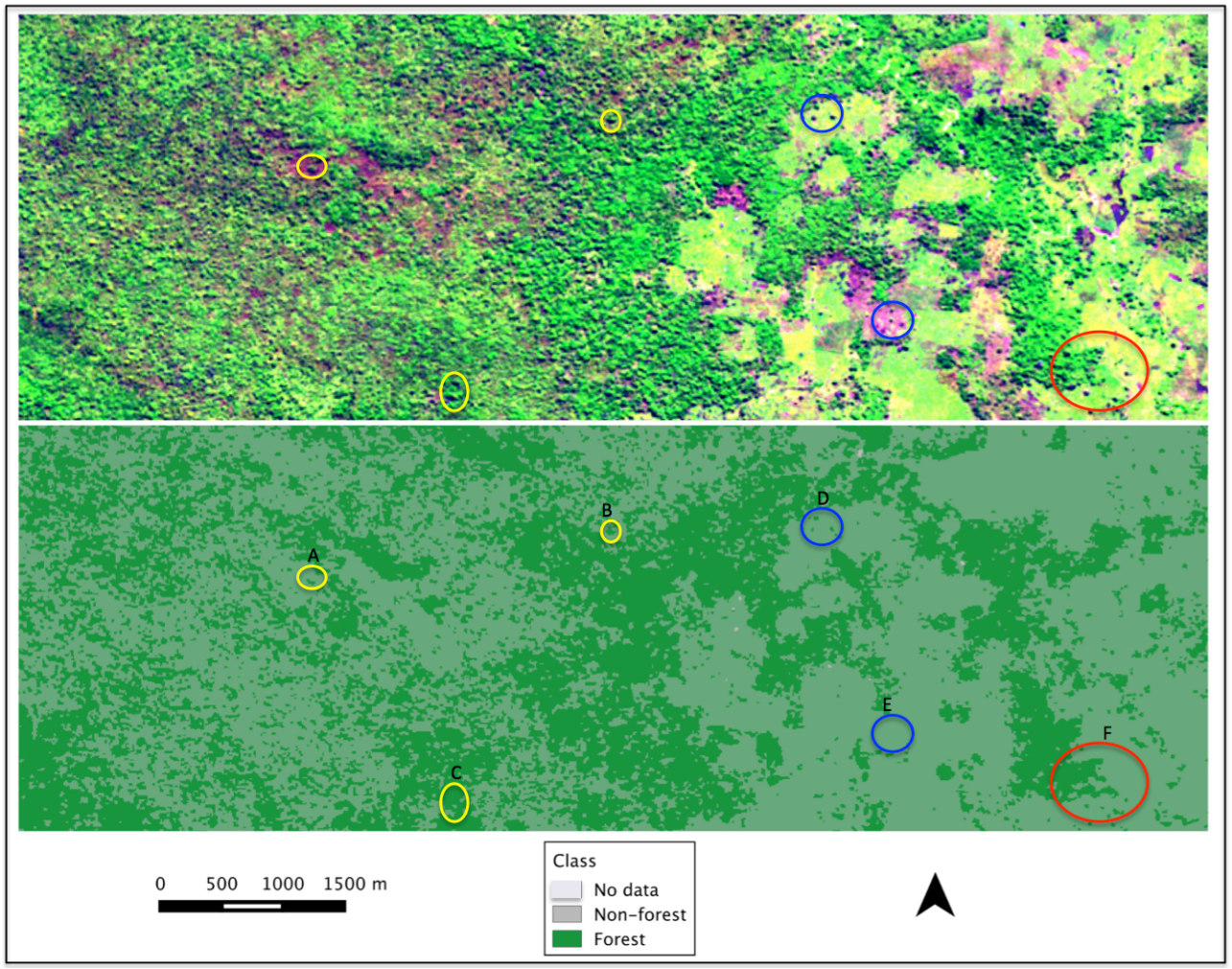


Figure 5.15: An illustration of parts of the classification that were possible sources of error. RapidEye false colour image was displayed as B:RedEdge, G:NIR, R:Blue.

5.4.2 Validation of the Classification

The importance of assessing the quality of thematic maps produced from remotely sensed data cannot be overemphasised. According to Foody (2002), this is a critical step for understanding error and its likely implications, and especially if it propagates through analyses linking the map to other datasets. In this work, the forest baseline thematic map would be used in the change detection, and therefore, errors being propagated in the process. Although it may appear simple in concept, accuracy is a difficult property to measure and express (Foody, 2002).

Foody (2002) summarised the problems associated with accuracy assessment, to highlight that even for basic approaches such as random sampling, there is still a suite of problems likely to arise. For Liwale, the process of validating the process was a challenge mainly because of the difficulties of sampling, and especially the shortage of suitable high high quality reference data. With those limitations, two validation approaches were undertaken for Liwale, and these being the comparison of the derived baseline product to LiDAR data of the site, and validation using Google Earth imagery. Measures that were derived from the confusion matrices included the overall accuracy and the kappa coefficient of agreement.

However, Olofsson et al. (2013) states that even though the overall accuracy and kappa coefficient measures are commonly published, these are insufficient to fully address the area estimation and uncertainty information needs. Moreover, Olofsson et al. (2014) recommended that the kappa coefficient be not used, especially for land use maps assessment. The study further lists limitations of this statistical measure, and these include that chance agreement is effectively overestimated in the calculation of the kappa coefficient, which in turn results in an underestimation of classification accuracy. Moreover, Foody (2002) argued that as a non-probability-based measure, the kappa coefficient is an inappropriate basis for accuracy assessment.

Foody and Mathur (2004) also stressed the need for more statistical measures to accompany accuracy statements. The study recommended that accuracy statements be accompanied by confidence limits, and that they be compared in a statistically rigorous fashion. This is because an accuracy statement is based on the class allocations observed for a sample of sites drawn from the map, thus only providing an estimate of the map's accuracy. A statistically-compared assessment, therefore, provides a more objective basis for comment and interpretation.

This research, therefore, further reported the accuracies in terms of 95% confidence limits and

overall accuracies that were adjusted for area as outlined in Olofsson et al. (2013). And in line with the recommendation of Olofsson et al. (2013) and Olofsson et al. (2014), the kappa coefficient was only used as a supporting statistic. The following subsections details out how these were undertaken for the two validation approaches that were undertaken.

5.4.3 Comparison of Forest Baseline Derived from RapidEye to LiDAR Data

The CHM derived from the LiDAR data flown over Liwale in February 2012 was used to undertake a comparison with the produced forest baseline in the area. Two comparison approaches were undertaken, first pixel-based validation (5 m), then over different grids (30 m^2, 100 m^2 and 250 m^2 grids). The choice of the 30 m^2 grid was motivated by that the methodology would hopefully be scaled out to Landsat, whose pixel is 30 m^2. Furthermore, the baseline was informed by the REDD+ forest definition, which defines a forest by percentage minimum height, canopy crown cover, and per hectare (100 m x 100 m), and thus the comparing of the classification to the CHM over a 100 m^2. Lastly, the comparison over a 250 m^2 grid had been motivated by that as part of the work, the aim was establish forest disturbance in Liwale, and then use the product together with a 250 m biomass map of the area to estimate biomass loss.

Systematic sampling approach was used in the comparison as the strips covered by the flight lines were used as the sample during the assessment. Figure 5.16 shows the total coverage of the forest baseline, and the LiDAR strips coverage area, which was used as the sample data for the comparison. So to sample, first, the baseline classification and the LiDAR forest/non-forest products were masked for the same spatial area as the LiDAR data (flight line strips), with areas between the strips assigned a 'no data' value (Figure 5.16).

The masked images were converted to binary images, Table 5.6 and Table 5.7 shows the values assigned to each class/category for images that were used to validate the classification using pixel-based and the grid-based comparisons respectively. Thereafter, the images were populated with statistics using the 'populate statistics' function in RSGISLib.

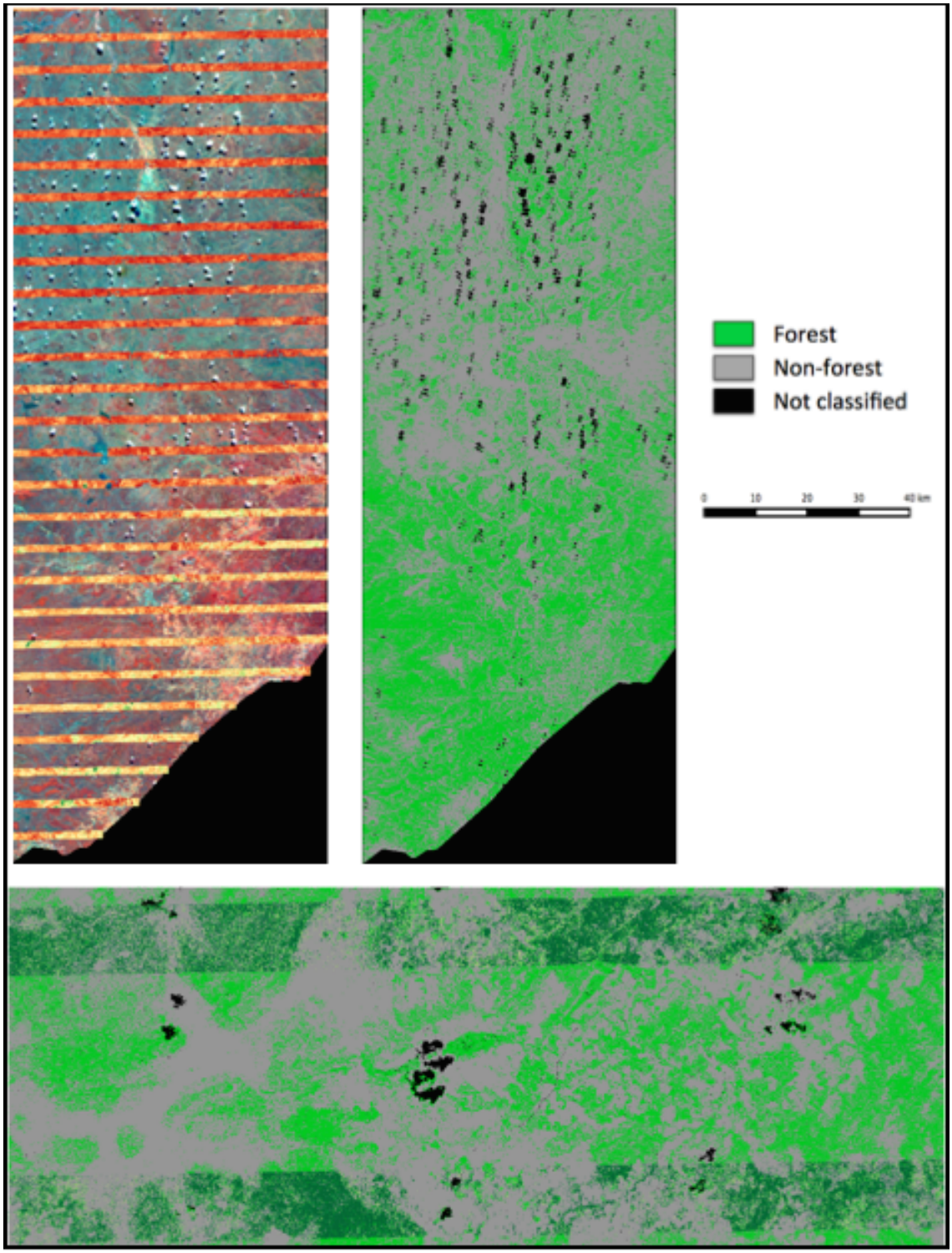


Figure 5.16: LiDAR strips used to compare the forest baseline classification and the CHM.

Table 5.6: The binary values assigned to classes in the RapidEye and LiDAR forest baseline products before the pixel-based comparison process.

Category	RapidEye classification	LiDAR	New assigned binary number
No data	0	-	999
Cloud/shadow	1	-	999
Non-forest	2	0	1
Forest	3	1	2

Table 5.7: The binary values assigned to classes in the RapidEye and LiDAR forest baseline products before the grid-based comparison process.

Category	RE classification	New RE binary values	LiDAR	New LiDAR binary value
No data	0	999	-	
Cloud/shadow	1	999	-	
Non-forest	2	1	0	10
Forest	3	2	1	20

Tanzanian Adopted REDD+ Forest Definition

To perform this comparison, the relevant CHM meeting the forest definition of Tanzania had to be used. As discussed earlier, for REDD+ monitoring in Tanzania, the adopted REDD national forest definition is an area with minimum tree crown cover of 10 %, minimum land area of 0.05 ha, and a minimum tree height of 2 m. As Table 5.8 shows, Tanzania is the only nation that has gone for the minimum permissible forest properties compared to other REDD countries, with most countries using the maximum value in the range of the three forest properties.

Table 5.8: REDD+ countries adopted forest definitions.

Country	Minimum tree crown cover (%)	Minimum land area (ha)	Minimum tree height (m)
Tanzania	*10*	*0.05*	*2*
Congo	30	0.5	3
Bolivia	30	0.5	4
Argentina	22.5	1	3
Cambodia	10	0.5	5
Equador	30	1	5
Indonesia	30	0.05	5
Panama	30	1	5
Paraguay	25	0.5	5
Viet Nam	30	0.5	3

Other than Tanzania, only Cambodia uses a minimum 10% crown cover, with the rest of the nations adopting a crown cover greater than 20%. It is worth noting that even though Cambodia

adopted a low minimum tree crown cover value, it uses a minimum land area of 0.5 ha and the minimum tree height of 5 m. Same with Indonesia, even though it adopted a low minimum land area of 0.05 ha, it uses the maximum values in the ranges of 30% minimum tree crown cover and 5 m minimum tree height.

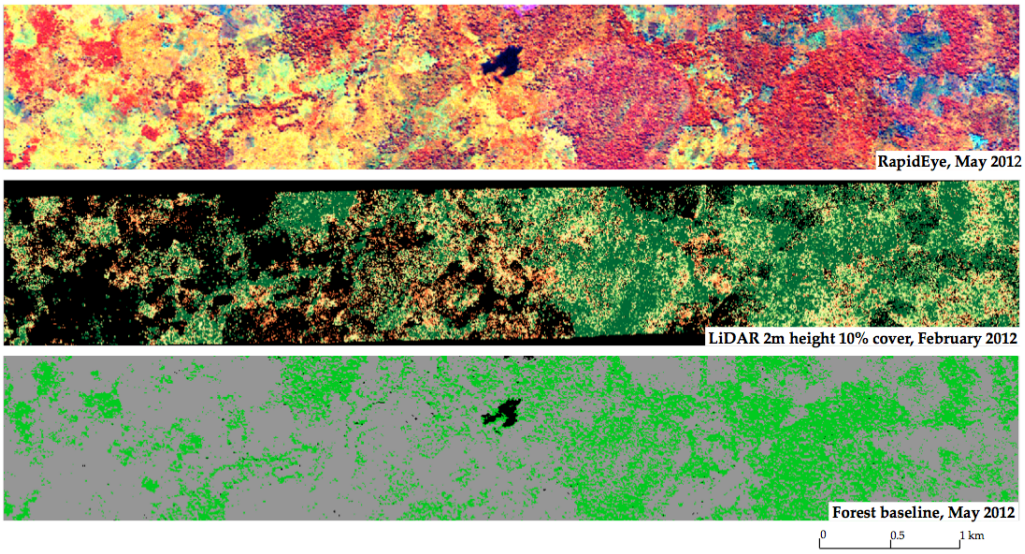


Figure 5.17: A subset of Liwale (SLU site) for the produced forest baseline, RapidEye false colour image (B:NIR, G:RedEdge, R:Blue), and the CHM of the currently adopted forest definition of Tanzania.

The forest definition is a major determinant of the kind of sensor data that can be used to map and monitor the forest resources of a country, and the level to which it can be monitored. This is because different sensors capture data at different parts of the electromagnetic spectrum, and therefore not all sensors may be appropriate for a certain environment. The adopted minimum forest definition variables have an implication on what can be achieved, and to what extent remote sensing data can be used for forest monitoring in the country, especially optical remote sensing data. Therefore, to compare the effect of the adopted forest definition of Tanzania, the CHM of Liwale meeting the minimum values of tree crown of 10 % and a minimum tree height of 2 m was compared with other canopy height models. All these models were then compared with the produced forest baseline of Liwale (Figure 5.18), a step that was undertaken to show how efficient and adequate the current forest definition was for establishing a forest baseline in the country using optical data, specifically, high resolution RapidEye data.

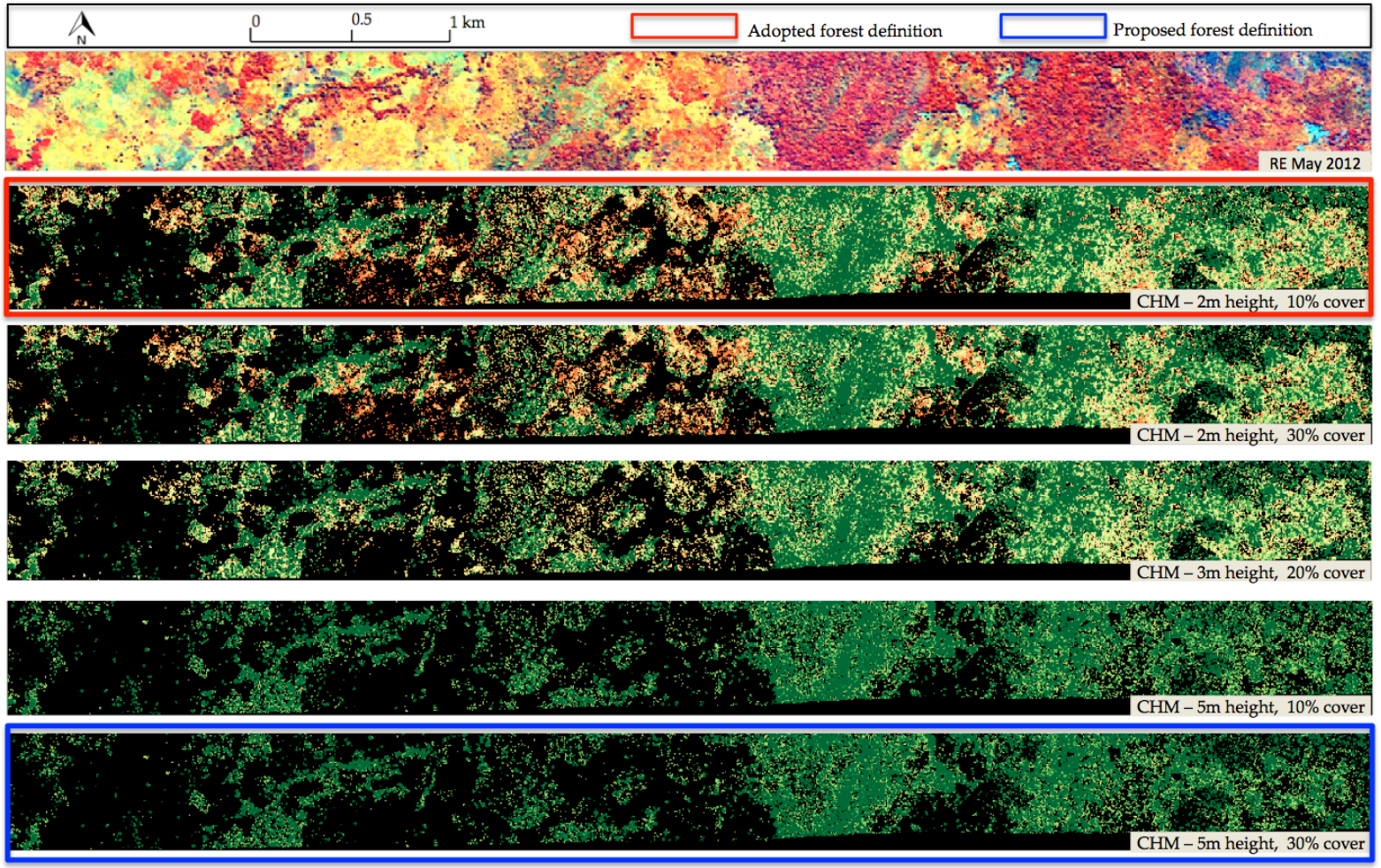


Figure 5.18: Different CHMs reflecting the effect of canopy height on the estimation of forest from optical satellite images.

An alternative approach that could have been used to get around the limitations introduced by the currently adopted forest baseline definition is the derivation of Foliage Projective Cover (FPC). However, Armston (2009) concluded that the relationships used to derive the FPC needed a very large amount of field data, and require the time of image data acquisitions to be reliable and transferable between image at different dates.

5 m Grid-based Comparison

In RSGISLib, the image math function was used to add two binary images, and Table 5.9 shows the resultant codes and errors. Thereafter, the pixel statistics were outputted for each of the four data types, and the overall accuracy derived.

Table 5.9: The codes and errors resulting from the addition of the binary images derived from the RapidEye forest baseline and the LiDAR CHM.

LiDAR class	**RapidEye class**	**New LiDAR binary**	**New RapidEye binary**	**Output**	**Error code/type**
Forest	Forest	20	2	***22***	True positive
Forest	Non-forest	20	1	***21***	False negative
Non-forest	Forest	10	2	***12***	False positive
Non-forest	Non-forest	10	1	***11***	True negative
Forest	Cloud/shadow	20	999	1019	-
Non-forest	Cloud/shadow	10	999	1009	-
No data	No data	999	999	1998	-

30 m, 100 m and 250 m Grid-based Comparisons

Furthermore, three grids (30 m, 100 m and 250 m) were also used to compare the derived forest baseline to the LiDAR forest product derived from the CHM of the proposed forest definition for Tanzanian REDD programme. The choice of a 30 m grid was motivated by the pixel size of a Landsat image. Since the proposed action plan was to further trial the methods developed for RapidEye using Landsat data, checking for comparability between the RapidEye forest baseline product and the LiDAR data, which would then be applicable to the Landsat forest baseline product as they would both be 30 m $\times$ 30 m. To further qualify that the overall accuracy was affected by the unsystematic shift in the RapidEye data, 100 m and 250 m grids were used. The motivation behind using the 100 m grid is that for REDD reporting, a hectare is used. A 250 m grid was used because it is equivalent to a carbon map product for Liwale, which was used to estimate carbon loss over the site between 2012 and 2014.

5.4.4 Comparison Results

For both the 5 m grid-based (pixel level) comparison and the other three grid-based comparisons, the CHM meeting the adopted REDD forest definition of the country (minimum tree height of 2 m and minimum canopy cover of 10%) was compared against the CHM that was found to be the most optimum for comparison with the forest baseline derived from RapidEye optical data (minimum tree height of 5 m and minimum canopy cover of 30%). Below are error matrices of the comparisons and statistic bias.

5 m Grid-based Comparison

At 5 m grid (pixel level), it was found that the results where close in terms of overall accuracy of the comparison of the forest baseline to the two CHMs. The CHM meeting the adopted forest baseline gave an overall accuracy of 71.8%, which was 1% better than the 70.8% for the optimum forest baseline, as summarised in Table 5.10 and Table 5.11 respectively. However, to be highlighted from these two comparison results is that it was found that when using the minimal forest definition values CHM tended to overestimate the forest class, which may be attributed to more pixels covered mainly by shrubs being classified. On the contrary, the CHM meeting the proposed forest definition for use with remote sensing data was found to underestimate the forest class. However, as discussed later in the chapter, the under-classification may have been as a result of a number of potential sources of error.

In terms of bias, it can be seen from Table 5.10 and Table 5.11 that there was large bias in the comparison of the produced forest baseline and the LiDAR CHM meeting the properties of the currently adopted Tanzanian national forest definition. However, the bias is less (reduced) in the comparison of the produced baseline to the LiDAR CHM meeting the properties of the forest definition being proposed for REDD monitoring. This shows that even though in terms of overall accuracy, the CHM product from the proposed forest definition has a lower percentage, it is still the better suited forest definition for use with remote sensing data that the one that the country currently uses.

Table 5.10: A matrix of pixel level comparison between the forest baseline and LiDAR CHM meeting the properties of the currently adopted country forest definition for Tanzanian REDD.

	Forest	*Non-forest*	*User*	**User (%)**
Forest	3459205	839762	4298967	80.47
Non-forest	1594746	2740613	4335359	63.22
Producer	5053951	3580375	8634326	
Producer (%)	68.45	76.55		**71.8**

Table 5.11: A matrix of pixel level comparison between the forest baseline and LiDAR CHM meeting the properties of the forest definition being proposed for the Tanzanian REDD.

	Forest	*Non-forest*	*User*	**User (%)**
Forest	2815972	1482995	4298967	65.50
Non-forest	1040218	3295141	4335359	76.01
Producer	3856190	4778136	8634326	
Producer (%)	73.02	68.96		**70.8**

30 m, 100 m and 250 m Grid-based Comparison

For comparison based on the 30 m, 100 m and 250 m grids, the comparison results (matrices) are summarised in Table 5.12. The summary include comparison of the forest baseline with the adopted REDD forest baseline, the proposed forest baseline, and other CHMs, a step taken to understand and account for how the properties used for forest definition and those adopted by Tanzania for REDD monitoring affect the use of remote sensing in forest monitoring.

The results show that for the two CHMs with the minimum three height of 2m, mapping forest cover using optical data leads to the mapping of a lot of non-forest areas into forest, and thus an over-estimation of the forest class. The landscape features that get picked up by the lower minimum height values, and thus lead to the over-estimation, are the areas vegetated with shrubs. Because of the over-estimation, the forest class had very high producer accuracies of over 86% across the three grids. But for the same reason, the non-forest class had a very low producer accuracy for these two CHMs across all three grids compared against. It was also found that the accuracy of the comparison between the two products got lower with an increase on the grid being used for these two CHMs.

A CHM that lies in-between the CHM of the adopted forest definition and the proposed forest definition, with a minimum tree height of 3.5 m and forest cover of 20%, produced an overall accuracy of about 71% for all three grids. The over-estimation improved as well, with the producer

accuracies for the non-forest class going to over 55% for all three grids, but still not yielding satisfactory results for forest monitoring at a national scale. The two CHMs with a minimum forest height of 5m in Table 5.12 showed an improvement on both the overall accuracies of the comparisons and the producer and user accuracies of both the forest and non-forest class. The CHM with a minimum forest cover of 10% gave an overall accuracy of 78% across the three grids, but still yielded slightly lower producer accuracies of less than 70%. But the CHM with a minimum forest cover of 30% produced overall accuracies of 79% or better for the three grids, and improved user accuracies for the non-forest class of about 70%. This is the CHM that gives the best comparison results to the forest baseline of Liwale.

Overall, across all the three grids employed in the comparison, it was found that at the present adopted forest definition for REDD in Tanzania, there was an overestimation of the forest class and an underestimation of the non-forest class.

For comparison at the 30 m grid scale, even though the overall accuracy for the comparison of the forest baseline product with the adopted forest definition CHM was 67.95%, the producer accuracy for the non-forest class was a low 41.38% and a relatively low user accuracy of 58.96% for the forest class. This underlines the scale at which the current forest definition leads to the overestimation of the forest class, but the underestimation of the non-forest class in optical remote sensing data. On the contrary, the CHM for the proposed forest definition yielded an overall accuracy of 78.87%, and the user and producer accuracies in this comparison were all over 71% for both the forest and non-forest classes.

Table 5.12: Summary of the different CHMs compared with the forest baseline of Liwale.

CHM	**Cover class**	**30m Grid**			**100m Grid**			**250m Grid**		
		Producer (%)	*User (%)*	*Overall (%)*	*Producer (%)*	*User (%)*	*Overall (%)*	*Producer (%)*	*User (%)*	*Overall (%)*
2m height & 10% cover	Forest	99.26	58.96	67.95	99.89	58.17	66.06	100	56.50	63.41
	Non-forest	41.38	98.51		35.84	99.73		30.28	100	
2m height & 30% cover	Forest	86.41	57.26	66.41	91.80	56.22	65.7	93.49	54.20	64.26
	Non-forest	51.31	83.35		45.99	88.13		43.28	90.26	
3.5m height & 20% cover	Forest	85.10	61.48	70.68	90.70	61.00	71.07	93.0	58.46	70.41
	Non-forest	59.80	84.19		56.26	88.91		54.99	92.0	
5m height & 10% cover	Forest	91.26	70.37	78.37	95.32	70.28	78.83	97.08	69.19	78.22
	Non-forest	67.45	90.11		64.19	93.91		61.38	95.92	
5m height & 30% cover	Forest	84.44	79.05	78.87	88.29	80.88	80.93	90.69	81.73	82.34
	Non-forest	71.87	78.62		70.56	81.03		69.99	83.55	

A similar pattern was observed for the comparison over a grid of 100 m. The overall accuracy for the comparison of the forest baseline product with the adopted forest definition CHM was 66.06%, but the producer accuracy for the non-forest class was a very low 35.84%, with a relatively low user accuracy for the forest class of 58.17%. This again shows the big the scale at which the current forest definition leads to the overestimation of the forest class and the underestimation of the non-forest class in optical remote sensing data. Again, the comparison of the baseline to the CHM for the proposed forest definition produced a much improved overall accuracy of 80.93%, and the user and producer accuracies over 71% for both forest and non-forest.

And again, the observed conclusions for the 30 m and 100 m grids comparison was found to hold true at the scale of a 250 m grid, with a slightly even higher overestimation and underestimation of the forest and the non-forest classes respectively. The overall accuracy for the comparison of the forest baseline product with the adopted forest definition CHM was even lower, just 63.41%. The producer accuracy for the non-forest class was also even lower, just 30.28%, and the user accuracy for the forest class was 56.50%. This still shows for remote sensing-based forest monitoring, the currently adopted forest definition may be inappropriate and incapable of meeting the forest monitoring needs. Again, on the contrary, the comparison of the baseline to the CHM for the proposed forest definition produced a much improved overall accuracy of 82.34%, with both user and producer accuracies of 70% and above for both forest and non-forest. Table 5.13 and Table 5.14 are the full matrices for comparisons performed over a 250 m grid using CHMs derived from the adopted forest definition and the proposed forest definitions respectively.

Table 5.13: Comparison matrix between the forest baseline and the LiDAR CHM meeting the currently adopted country forest definition for Tanzanian REDD over a 250 m grid.

	Forest	*Non-forest*	*User*	**User (%)**
Forest	861	663	1524	56.50
Non-forest	0	288	288	100
Producer	861	951	1812	
Producer (%)	100	30.28		**63.41**

Table 5.14: Comparison matrix between the forest baseline and the LiDAR CHM meeting the proposed forest definition for Tanzanian REDD over a 250 m grid.

	Forest	*Non-forest*	*User*	**User (%)**
Forest	984	220	1204	81.73
Non-forest	101	513	614	83.55
Producer	1085	733	1818	
Producer (%)	90.69	69.99		**82.34**

Similarly, as was evident with the pixel-based comparison, it was noted that at grid-based comparison, there was generally large bias for products from the currently adopted forest definition, and less bias for products from the forest definition being proposed. Table 5.13 and Table 5.14 show such for the 250 m grid comparison. Besides having a much higher overall accuracy of 82% (compared to the 63% for the comparison of the baseline to the CHM product from the currently adopted forest definition), Table 5.14 which is the comparison of the baseline to the CHM product from the proposed forest definition also has less bias compared to Table 5.13. This consistency for all three grid levels shows that the CHM product from the proposed forest definition is better suited for use with remote sensing data that the one that the country has adopted.

5.4.5 Comparison of Forest Baseline to LiDAR CHM Using Scatter Plots

To further compare the forest baseline to the CHM of the proposed REDD forest definition, as found to yield the most accurate forest cover estimates, the two products were plotted against each other for 30 m, 100 m and 250 m grids, and scatter plots were produced. Figure 5.19 shows the plots for the three grids respectively. The plots depict that, even though there are features/polygons incorrectly classified, which may be due to different possible sources of error in the classification, most of the data falls along the linear line. The data notably got even more concentrated to the linear line with an increase on the grid size. However, this may just be as a result of fewer resultant total number of counts for the bigger grid than the smaller ones since the R^2 value of all three grids is relatively good, and all had a confidence value of 99%. Table 5.15 shows that the 30m grid had an R-square value of 0.75, while both the 100m and 250m grids had a value of 0.81.

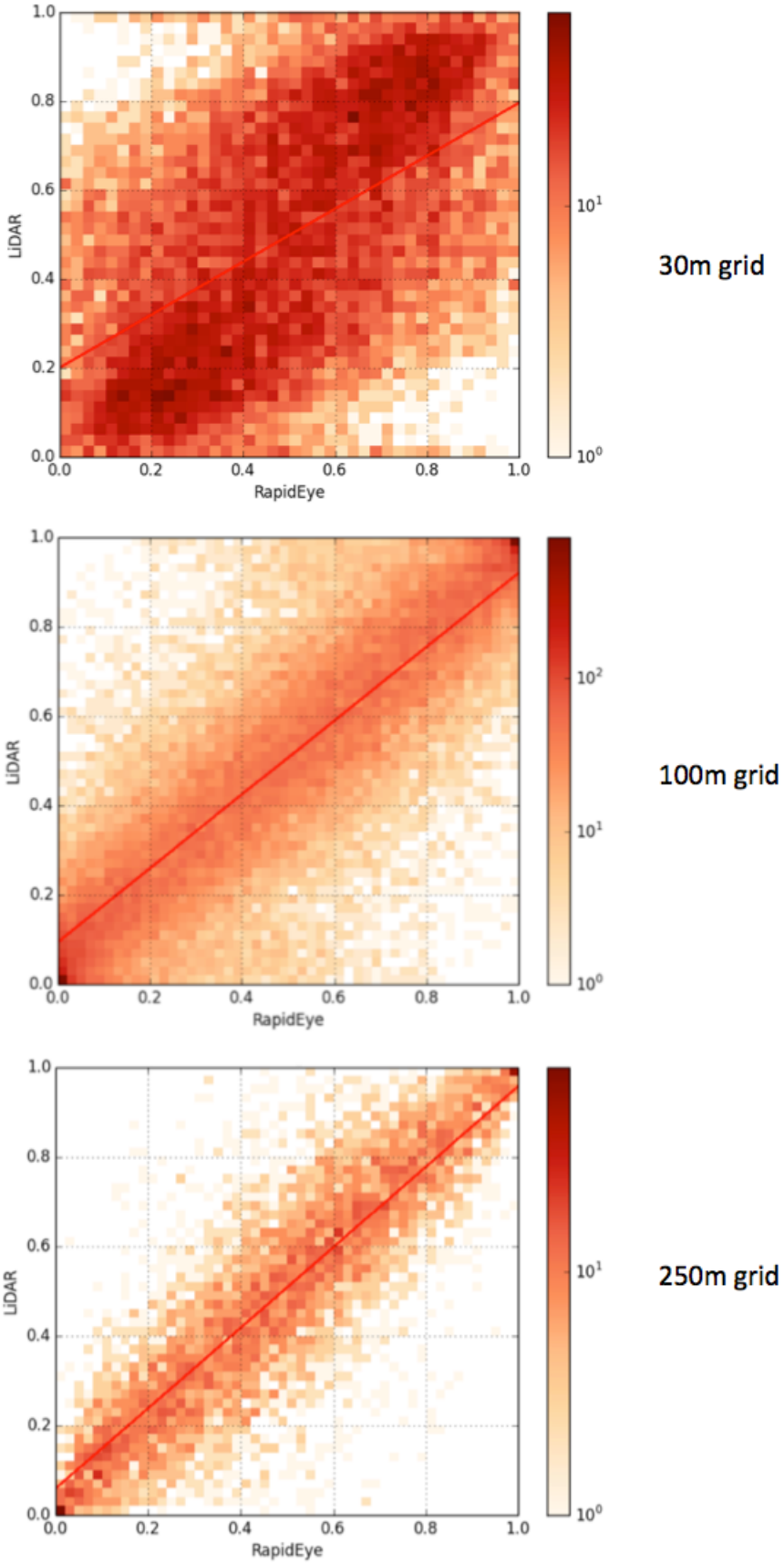


Figure 5.19: Scatter plots of the comparisons of the forest baseline and the LiDAR CHM of the proposed REDD forest definition.

Table 5.15: The R-squared values of the comparison between the products over the different grids and their confidence values.

Comparison grid	R^2	Confidence (%)
30 m	0.746	99
100 m	0.808	99
250 m	0.816	99

5.4.6 Validation of Forest Baseline Using Random Points

As mentioned, there was no field data for Liwale that was available for use for validation purposes in the study. Moreover, the LiDAR data, even though assumed to be accurate, had not been validated, and thus its use only at a comparison level. A commonly used approach in literature where there was no or inadequate field data for validation is the collection of ground truth data by visual interpretation of high resolution optical data available through Google Earth (Schneider et al., 2009; Balthazar et al., 2012; FAO and JRC, 2012; Clewley et al., 2014; Dons et al., 2016). To further assess the accuracy of the classification, the study adopted a similar approach, where Google Earth imagery and RapidEye data as the reference datasets. The land cover or land use type for sample points was visually corrected for each point, and these used to validate the baseline classification.

Ideally, the research should have just used the Google Earth imagery as the primary reference image. However, the 'original' RapidEye data was also used in the accuracy assessment as means to avoid introducing errors into the reference dataset for temporally sensitive classes (Kahya et al., 2010) or poor coverage of Google Earth imagery over some parts of the area. Poor coverage would be that an older image was available, a seasonally inappropriate image (dry season image) or there was no data due to cloud cover or cloud shadow.

It was found that some areas within Liwale had very poor coverage in Google Earth. Noted limitations with Google Earth imagery included that some areas have no cloud-free imagery (Figure 5.20), some have very poor quality images. In addition, for some areas, the acquisition time was different between the Google Earth image and the RapidEye image that was used to produce the forest baseline. Besides these challenges of being limited due to dates and seasonality, it was noted that some images making up the Google Earth coverage over Liwale were poorly registered. The Google Earth coverage is a combination of the best available image of different sensors, which unfortunately are not registered to the same level of accuracy. Over Liwale, there was quite a

large coverage covered by RapidEye, which unfortunately, as already highlighted, is mainly dry season images and generally have a poor registration. The registration issues for RapidEye is more pronounced in Africa and South America, which is attributed to poor availability of GCPs in that part of the world.

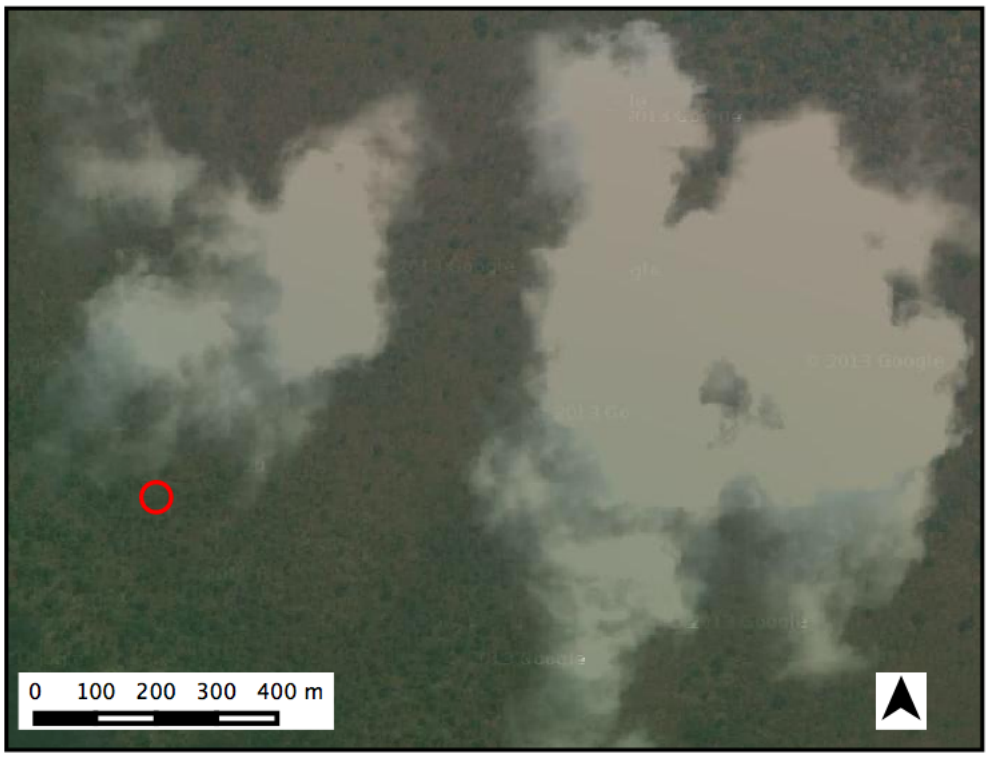


Figure 5.20: An illustration of an area in Liwale whose best Google Earth image was covered with cloud, which would mean that area is without data for validation.

Figure 5.21 is an example of two different sensor images with a reasonable image registration difference between them, as evidenced by the stretch of road cutting across them. Unfortunately, for these heavily fragmented savanna landscape of Liwale, such a poor image registration has a potential to negatively influence the validation to a large scale, even if it is off by a few metres. These noted limitations with the Google Earth dataset therefore forced the use of the RapidEye false colour composite image as the primary reference image, and Google Earth image used as the secondary image.

Figure 5.21: An illustration of poor image registration on the Google Earth image of Liwale that was meant to be used as the primary validation image.

Sampling Approach

The sampling approach used was simple random sampling (SRS), which was undertaken as a GIS procedure, whereby points for each class were randomly selected, and then validated if those land areas had been correctly classified for through visual interpretation of the high resolution reference

data. SRS assumes that the population is independent and identically distributed, with the sample chosen from the population independently with equal probability, and inferences conducted using that sample (Chander et al., 2009).

Systematic random sampling (SYS) is known to produce more precise estimates than SRS (Aune-Lundberg and Strand, 2014). However, SRS was preferred in this study. Its preference was motivated by that, potential disadvantages of SYS include that it can lead to poor precision if the sampling interval happens to coincide with periodicity in the population. An example is if in a product, the classification error is spatially periodic (Stehman, 2009). Moreover, Olofsson et al. (2014) noted that it was essential that only unbiased or consistent estimators be used. However, studies such as Stehman (2009) and Aune-Lundberg and Strand (2014) concluded that it was not possible to construct an unbiased estimator of variance for SYS in land use and cover studies.

Olofsson et al. (2013) also recommended the use of SRS compared to other sampling approaches for validating land use and cover, and change maps. A study by Congalton (1991) performed sampling simulations on three spatially diverse areas and concluded that in all the cases SRS produced similarly satisfactory results as SYS. Even though it is sometimes thought to be a simple sampling approach, the approach very effective, and thus adopted in this study.

Selection of Sample Points

The sampling process involved first generating a shapefile of random points from the produced 2012 forest baseline. This was done through the use of a function in the classification module of RSGISLib called *generateRandomAccuracyPts*. Having assigned the forest baseline as the input image, each generated point was populated with its cover class (forest or non-forest) in the attribute table of the resultant shapefile.

Then in QGIS software, a 5 m buffer was created around the sampled points. Thereafter, an accuracy assessment tool called Class Accuracy (Bunting, 2015) that had been installed as a plug-in in QGIS was used to validate the sampled points. Validation in the tool was done through comparing the assigned class in the classification for each buffered point against the reference data (RapidEye and Google Earth images). Using a 5m buffer meant that a window of four pixels was considered for each sample point, which was more representative of the environment around, than considering just a single pixel point. This was more especially useful since the RapidEye data was poorly registered, and thus this was a step that would compensate for and/or reduce errors arising

from the poor registration of the image.

The accuracy assessment tool enables a user to zoom to one sampled point at a time during the validation process, as illustrated for the buffered sample point (red circle) in Figure 5.22. In this example, the selected point had been classified as a forest, which is the information it had on the attribute table. Having loaded the reference images (RapidEye and Google Earth) into QGIS, the image was used to confirm whether each of the sampled points had been correctly classified. The reference images (Google Earth image on the illustration) shows that it was correctly classified as the point falls in a forested area.

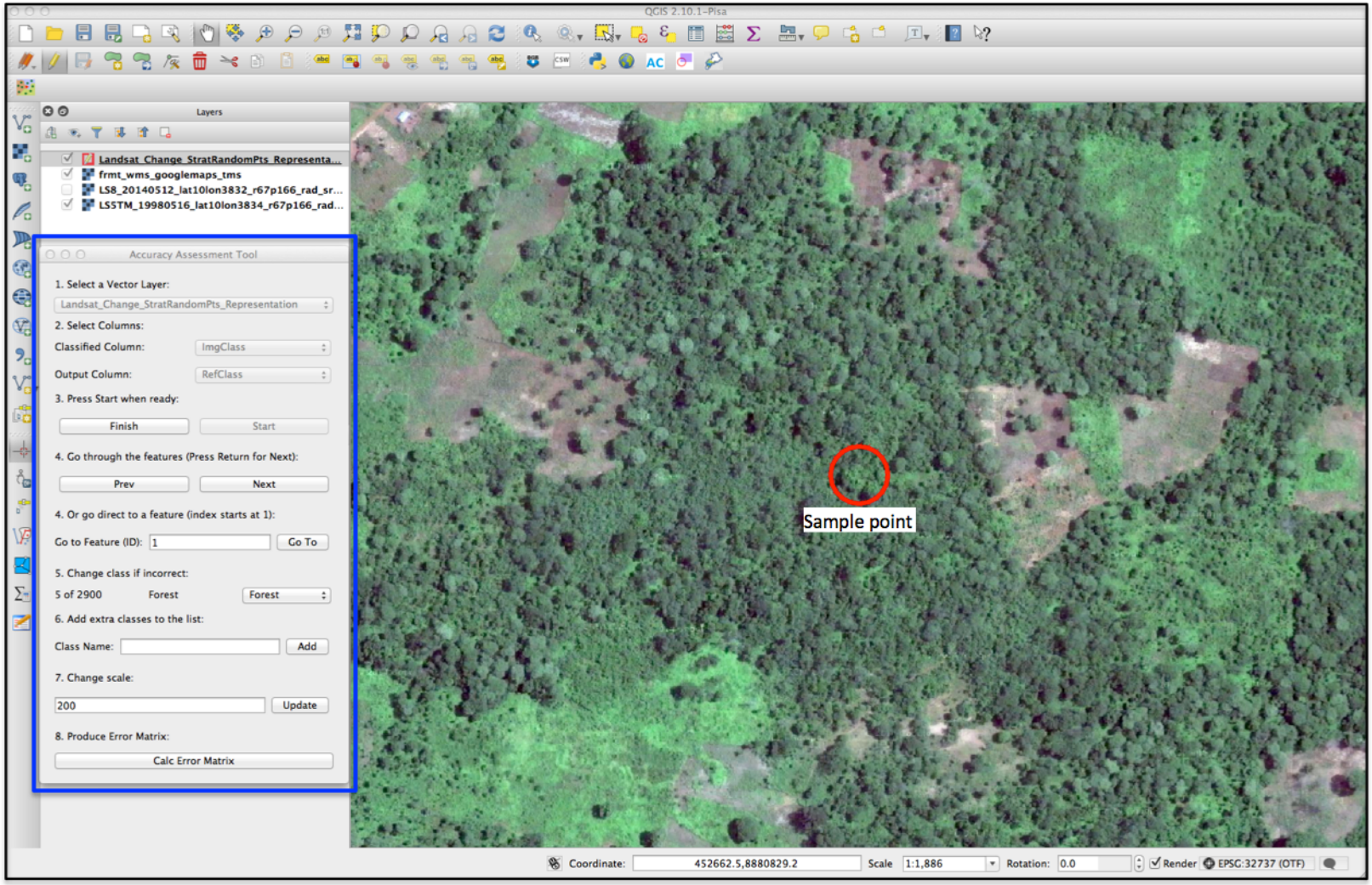


Figure 5.22: The accuracy assessment tool used as a plug-in in QGIS.

For sample points that had been correctly classified, the same classification class was confirmed in the reference column in the attribute table. But if the classification was found to be incorrectly classified over a sample point, a drop-down list was used to select and assign it the correct class. In the process, this was updated into the reference column of the attribute table. For new scenarios that were not accounted for in the drop-down list, that new class was added. On completing this step, an error matrix was outputted for each of these set of sample points, which yielded the user, producer and overall accuracies of the classification, as well as the kappa coefficient.

Validation Results

With Liwale site having been divided into northern and southern Liwale during the classification, the validation was undertaken independently, by randomly sampling, as well as sampling across the transects for each of the two parts of the site. For each part of the site, the two sets of samples were used to assess the classification accuracy, before an overall accuracy for that site was computed by combining the two sets of sample points. Table 5.16 is a summary of the overall accuracies and kappa coefficients from the random points, transect points and the combined points for northern and southern Liwale. The table also shows the overall accuracies and kappa coefficients for the whole site which was derived from combining the set of points from both parts of the site. The results showed consistency as overall accuracies relatively ranged between 95% and 97% for both northern and southern sites, as well as for the whole site.

Table 5.16: Summary of the error matrices for random points, transects points and the combined points, and their kappa coefficients.

Location	**Random points**		**Transect points**		**Combined points**	
	Overall accuracy	*Kappa*	*Overall accuracy*	*Kappa*	*Overall accuracy*	*Kappa*
Northern Liwale	95.25	0.872	97	0.944	94.87	0.881
Southern Liwale	96.15	0.924	94.21	0.883	96.67	0.929
Whole site	95.61	0.903	96.02	0.921	95.81	0.914

Below are the error matrices for validation across the whole site, using simple randomly selected points, points sampled across transects, and combined points from both SRS and transects. Table 5.17 is the matrix for the randomly selected points throughout the site, and they were 2210 in total. It shows that the forest baseline product had been classified to an overall accuracy of 95.61%, with a kappa coefficient of 0.903. The table shows that both the forest and non-forest classes were

well classified for as the both classes had producer and user accuracies of 93% or greater.

Table 5.17: Error matrix for 2210 randomly selected points within Liwale which yielded an overall accuracy of over 95%.

	Non-forest	*Forest*	*User*	**User (%)**
Non-forest	1363	60	1423	95.78
Forest	37	750	787	95.3
Producer	1400	810	2113	
Producer (%)	97.36	92.59		**95.61**

The same was done for the set of points from the transects, and Table 5.18 is the matrix for the transects points, which as stated, were on the boundaries of the two classes. In total, they were a sample of 2160 points. The matrix shows that around class boundaries, the overall accuracy was slightly better than generally across the image (random points), with an overall of 96.02%, with a kappa coefficient of 0.921. Again, both the forest and non-forest classes were well classified for as the both classes had producer and user accuracies greater than 94%.

Table 5.18: Error matrix for 2160 points that were generated across boundary transects in Liwale.

	Non-forest	*Forest*	*User*	**User (%)**
Non-forest	1052	27	1079	97.5
Forest	59	1022	1081	94.54
Producer	1111	1049	2074	
Producer (%)	94.69	97.43		**96.02**

To determine the overall accuracy of the forest baseline of Liwale, the two set of points were merged into one dataset (shapefile), which produced in total, 4370 sample points. For this new dataset, an error matrix produced, which gave the overall accuracy and kappa statistic of the classification. The overall accuracy of the produced baseline of Liwale was 95.81%, with a kappa coefficient of 0.914, which shows a relatively good forest/non-forest baseline was successfully produced for the savanna forests of Liwale using the LCCS-adapted EODHaM classification system.

Table 5.19: Error matrix of the forest baseline from RapidEye data all sample points of Liwale.

	Non-forest	*Forest*	*User*	**User (%)**
Non-forest	2415	87	2502	96.52
Forest	96	1772	1868	94.86
Producer	2511	1859	4187	
Producer (%)	96.18	95.32		**95.81**

For this combined random points dataset, accuracy measures were then further presented with 95% confidence interval, as recommended by (Olofsson et al., 2013). The forest baseline map had an overall accuracy of 0.94 ± 0.012. The user's accuracies were 97% for the forest class and 93% for the non-forest class, with uncertainties of ±0.8% and ±1.8% respectively. These accuracies were relatively similar to those derived from the use of an error matrix from the baseline map. However, for the producer's accuracies, it improved to 98% for the non-forest class, with an uncertainty of ±0.5%. And notably, for the forest class, the producer accuracy decreased to 88%, with an uncertainty of ±2.5%. Table 5.20 is a summary of the user's, producer's and overall accuracies presented with a 95% confidence interval.

Table 5.20: Error matrix for the 2012 forest baseline map of Liwale derived from RapidEye data. Map categories are the rows while the reference categories are the columns. The accuracy measures are presented with a 95% confidence interval and the overall accuracy was 0.94 ± 0.012.

Class	***Forest***	***Non-forest***	***Total***	***Am [sq. km]***	***Wh***	**User's**	**Producer's**
Forest	1363	37	1400	2088.94	0.366	0.97±0.008	0.88±0.025
Non-forest	60	750	810	3620.64	0.634	0.93±0.018	0.98±0.005
Total	1423	787	2210	5171	1		

5.5 Discussion

5.5.1 The LCCS Classification and the Forest Baseline

The EODHaM LCCS classification is a rule-based approach, using spectral dissimilarities to develop unique rule-sets to describe the classes of interest. A popular approach with most researchers of late has been the use of machine learning techniques. But their over-reliance on a training dataset makes them not suited for the African environment where there is usually insufficient or no field data. Lack of or insufficient ground reference (field data) normally result in unrepresentative training datasets (Kotsiantis, 2007), which in turn affect the results outputted. Even though it can be used to enhance the results if available, LCCS does not rely on field data as rule-sets can be developed only from the spectral bands of the image. This makes it well suited for Tanzania and the African environment where field data are lacking. The National Forestry Resource Monitoring and Assessment (NAFORMA) forest inventory dataset is not accessible, which would otherwise go a long way in availing ground reference data for remote sensing research in the country.

Robustness and transferability of the technique advanced for the environment was sought to ensure

it can is scalable to a national level and to other environments in the SADC region. Machine learning techniques were reported in literature to be relatively difficult to transfer and apply in other environments. The results from the technique developed using the LCCS lines up with other studies such as Drgu et al. (2014) and Lucas et al. (2015), that concluded that the LCCS classification produced scientifically acceptable results in LULC mapping. Even though these studies used the classification under different environments, this study concluded that it was effective for the mapping African savanna forests. The forest baseline was classified for with an overall accuracy of 95.81% and a kappa coefficient of 0.914. This is an accuracy and kappa coefficient amongst the best attained in literature (see Table 5.1). Studies using the rule-based approach for forest/vegetation mapping that obtained overall accuracies of 90% or better include Bardossy and Samaniego (2002), Kahya et al. (2010), Zhang and Zhu (2011), and Myint et al. (2011). Other studies reported overall accuracies less than 90% (Lucas et al., 2007; Lewinski and Bochenek, 2008; Tseng et al., 2008; Suchenwirth et al., 2012; Belgiu et al., 2014; Yang et al., 2015), which shows that the accuracy achieved was amongst the best for studies that used the rule-based approach.

The accuracy measures were further validated through the use of 95% confidence level statistical measure, which provided further information on uncertainty (Olofsson et al., 2013). Presenting the accuracy measures with a 95% confidence interval factored in the areal coverage of each class. The measure highlighted that there was actually much higher omission in the classification for forest in Liwale, with the producer's accuracy for the class being 88% (compared to 96% derived from the image-based error matrix), with an uncertainty of 2.5%. This was notably the highest uncertainty as the other user's and producer's accuracies had lower then 2% uncertainties. The producer's accuracy of the non-forest class was improved as it was 98% (compared to 96% derived from the image-based error matrix). The user's accuracies for both forest and forest classes were found to be consistent with those derived from the image-based error matrix. The statistical method proved to be very effective in providing more information on the accuracies than what error matrices can give.

Among the major limitations of earlier use of rule-based classifications was the challenge with only a few such studies successfully developing generalised rule-sets that are transferable to other environments. The aim of this research was to develop a forest mapping technique that would be transferable to other savanna environments in Tanzania and the African region, and the developed technique holds that potential.

The overall accuracy of the classification would have been even better with a good time-series data available in Liwale. However, seasonality and persistent cloud cover greatly reduced the availability of data appropriate for forest mapping. In years where there were more image acquisitions, as documented in Chapter 4, such data was found to have seasonal effect. Moreover, RapidEye data only dates back to the year 2010 in Tanzania. With a better time-series data over the Liwale area and the savanna environment generally, the potential for producing a forest baseline of high accuracy would be enhanced. Even though the data presented a lot of challenges, and the method had to be modified to accommodate data limitations, the forest baseline product had a high accuracy, and the conclusion was that with more data being available, the technique will guarantee forest mapping that yield a baseline of high quality.

Another challenge with the savanna forests of Liwale is their high level of heterogeneity. This leads to an overlap in the feature space, and thus complexity in the separation of the classes during the segmentation and classification process. During the segmentation stage, the minimal segment/feature size was reduced to only two pixels. However, this could not eliminate the production of some segments that combined different classes. Partly, this was also as a result of the resolution of RapidEye (5 m), that even though a high resolution imagery, it is a spectrally poor image that is not capable of showing the boundaries of the classes. Poor segments with overlaps to other classes would result in a classification with those overlaps incorrectly classified. The classes that had very similar spectral signatures also contributed to a confusion between these classes during the classification process. For example, wetland areas in the area were incorrectly classified as forest areas. Young and productive forest areas were found to confuse with fields with crops, grasslands and shrubs in some areas. The boundaries between these spectrally similar classes were also difficult to determine, even though the transect points between such classes showed that these regions were also well classified.

Up until recently where Black Bridge is working towards rectifying this challenge, RapidEye was poorly registered in Africa and South America due to limited ground control points in these regions (RapidEye, 2012). In this study, the limitation of the poor registration of the dataset was pronounced, and it was found that this was the case even between individual scenes within the Liwale site. The shift in the dataset was not systematic, which made it almost impossible to correct for it as this would have to be done per scene, with the user having to determine the direction and distance of the shift. Noted directions of the shift within the 11 scenes for the Liwale site include both the straight and diagonal directions. The extent of the poor registration of the

dataset was further noted on the Google Earth image where it was noted to be off compared to other datasets, as illustrated in Figure 5.21.

5.5.2 Adopted Forest Definition for REDD

Table 5.12 shows that the adopted REDD forest definition of a REDD country determines to what extend and what ability can remote sensing data adequately and effectively map the forest baseline of the savanna environments. As Magdon et al. (2014) alluded to, different forest definitions have different effects on the forest cover maps and estimates. In this proposed approach where the baseline is further used in the estimation of forest change (loss and gain), the limitation and the errors due to this would be propagated further into the monitoring system. For the range of forest definitions trialled, the results showed that the lower the forest height and the forest percentage cover used, the lower the ability of optical remote sensing data to adequately monitor the savanna forests of Tanzania.

One major conclusion from the comparison of the different CHMs to the derived forest baseline was that the height attribute was more crucial than percentage cover, even though percentage cover itself does affect the results. This findings highlighted the importance of the use of an appropriate forest definition to ensure the adequacy and effectiveness of remote sensing data for forest monitoring in the country. The REDD programme defines the range of minimal values a country may adopt to define a forest, and further encourages the REDD countries to adjust the country definition as per need to ensure that the monitoring process is a success. The REDD+ programme is built around the use of remote sensing data for mapping and monitoring the forest resources (UNFCCC, 2002). In Tanzania where the adopted forest definition was found to be inappropriate and incapable of adequately mapping and monitoring the forest resources, it may be advisable to consider modifying the current minimal forest properties.

5.5.3 Sources of Error in Classification and Validation

The major sources of error in the classification for the baseline in Liwale included that this savanna environment is complex and highly heterogeneous which leads to a lot of overlap in the feature space. Another challenge that was encountered was the shift in the RapidEye imagery, which was reported in other studies (e.g. Krystyna et al. (2009); Chander et al. (2013)). With the

environment highly fragmented, this shift is capable of introducing error to either class as features are off, and thus incorrectly classified. When a comparison is then done with a product such as the LiDAR data which had detail so as to model even features made up of just an individual or a couple of pixels, a shift may have a large effect. To compensate for the limitation, comparison was undertaken at grid level, than at feature level. However, error compensation works best for not heavily and heterogeneously fragmented environments.

The adopted validation method ensured that the shift in RapidEye data did not have a negative influence on the validation process. First, the validation process was performed against the image data, from which the results were derived. Therefore, the shift would not affect validation. Secondly, Google Earth image was further used together with the original image in the validation. In addition, sampled points were validated using a window of pixels, than a single point, which too was aimed at cancelling out the effect of any errors arising from the shift. It is believed that these steps ensured that the shift did not have a negative effect on the accuracy assessment.

Another source of error that was related to the data, and hard to eliminate for, was the difference in the acquisition times of the datasets that were used. In this case, the baseline and the CHMs it was compared to were produced from data that had been acquired three months apart. The RapidEye imagery that was used to produce the forest baseline was acquired in the month of May 2012, while the LiDAR mission was undertaken in February 2012.

Even though only three months apart and both of them within the wet season window that is appropriate for forest mapping and monitoring in Tanzania, it is worth noting that each was at the extreme end of the season. The LiDAR data was acquired at the beginning of the wet season, yet the RapidEye data was acquired at the end of this season. Within the three months, there had been substantial forest change within some areas, as evidenced in Figure 5.23. Circled in blue are areas that were forested during the month of February, when the LiDAR data was acquired. But by the end of the wet season (May), these areas had been cleared and converted to agricultural fields. This too introduced errors to the overall accuracy of the forest baseline product.

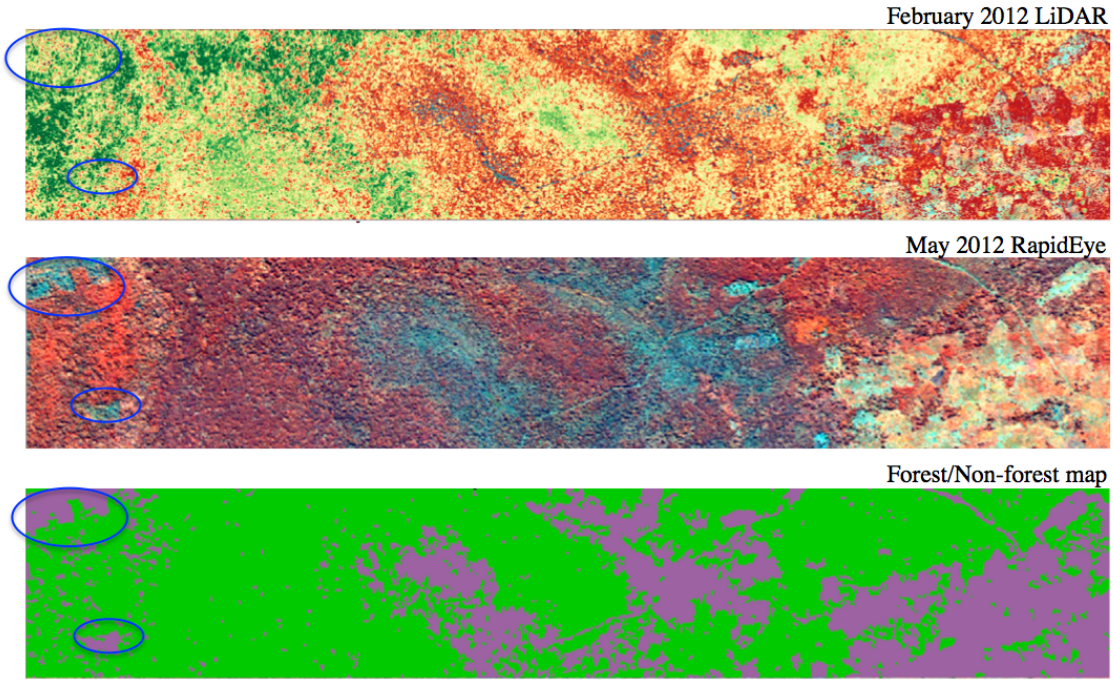


Figure 5.23: Circled in blue are areas that were forest during LiDAR acquisition in February 2012, but had been disturbed during RapidEye acquisition in May 2012.

5.6 Conclusions

A technique that uses a single date RapidEye imagery to map the forest baseline of the savanna woodlands of Liwale, southern Tanzania was successfully developed. This is technique found to map forest from non-forest with an overall accuracy of 95.81% and a kappa coefficient of 0.914 for the savanna forests of Liwale. When presented with 95% confidence interval, the overall accuracy was 94 $\pm$ 1.2%. The method has the capability to be transferred across to the wider savanna environment of Tanzania and the African savannas generally. The forest/non-forest product forms a crucial product that has the potential to serve as a baseline for future monitoring purposes. Data availability due to seasonality and cloud cover is the main hindrance to the availability of appropriate data for forest mapping in Liwale and Tanzania.

Chapter 6

Change Detection from RapidEye Imagery

6.1 Introduction

Disturbance refers to any relatively discrete event that disrupts the ecosystem, community, or population structure, leading to change in resource availability or the physical environment. It can occur due to natural hazards (such as natural fires, volcanoes and earthquake), or from anthropogenic processes (such as urban growth, deforestation and degradation). Disturbance tend to lead to the altering of the state and trajectory of an ecosystem, and thus are key drivers of spatio-temporal heterogeneity. This process contributes to the increased carbon dioxide emissions from terrestrial biomass loss (Verbesselt et al., 2012). According to Turner (2010), disturbance is a key component of ecological systems that affects numerous ecosystems across a wide range of scale, such as terrestrial, aquatic, and marine ecosystems. Many studies have developed techniques for accounting for disturbance through detecting change.

Changes in forest cover affect the delivery of important ecosystem services, including biodiversity richness, climate regulation, carbon storage, and water supplies (Hansen et al., 2013). Change detection is the process of determining and/or describing changes in LULC properties based on co-registered multi-temporal remote sensing data (Verbesselt et al., 2012). Singh (1989) defines it as the process of identifying differences in the state of an object or phenomenon through observing

it at different times. Quite a wide range of remote sensing-based change detection algorithms have been developed (Coppin et al., 2004). The basic premise in using remote sensing data for change detection is that the process can identify change between two or more dates that is uncharacteristic of normal variation (Verbesselt et al., 2012). In the last three decades, numerous researchers have addressed the problem of accurately monitoring LULC in a wide variety of environments (e.g. Coppin and Bauer (1996); Coppin et al. (2004); Lu et al. (2004); Podeh et al. (2009); Verbesselt et al. (2010, 2012)). Monitoring and detecting change, especially over a broad scale, provide a platform for deriving information on phenomena such as LULC, disaster monitoring and shifting agriculture patterns (e.g. Alqurashi and Kumar (2013)).

Having established the forest baseline for Liwale from the 2012 wet season RapidEye image, a technique for detecting potential change features, and for classifying these into either change or non-change features was developed. This chapter outlines how this was done and shares the results achieved when using a 2014 wet season RapidEye image to detect change. The approach uses a baseline and a new acquisition image to identify potential change features, then these are classified to produce a change map and an updated baseline for the new year of acquisition.

To investigate ecosystem disturbances using satellite data, many approaches focus on establishing if and where disturbances occur in an observed time series (Verbesselt et al., 2010, 2012). Remote sensing-based techniques are not only consistent and repeatable, but they permit for the efficient incorporation of data attributes even from the non-optical parts of the electromagnetic spectrum (Coppin et al., 2004). Another relatively new approach used by Verbesselt et al. (2012) investigated if new observations still conformed with the expected behaviour of a historical sample. This method therefore, detects disturbances at the end of a time series by comparison with representative, stable, historical observations. It is capable of detecting disturbances within newly acquired time series data by automatically identifying a stable history period, which in turn is used to model the season-trend variation, against which disturbances can be detected (Verbesselt et al., 2012).

Every remote sensing-based change detection approach is affected by either or all of the following limitations; spectral, spatial, temporal and thematic constraints. Reviews of change detection techniques such as Shoshany (2000), Coppin et al. (2004), and Podeh et al. (2009), highlighted that besides the standard problems encountered in change studies such as registration error, variation in atmospheric illumination and sensor variability, there are other potential sources of error. These include the effects of topography, heterogeneity in vegetation types and inter-annual phenological

variability resulting from variable precipitation patterns. Moreover, Coppin et al. (2004) stressed the importance of selecting the right method for implementation as each has the ability to differently influence both the qualitative and quantitative change estimates, which means that different approaches can yield different change maps for the same environment. It highlights the need for a careful consideration of the method being chosen to ensure it is well suited for the data type, environment being studied, and the change being detected. The following sections are a review of different commonly used algorithms and their properties (advantages and disadvantages).

6.2 Background

6.2.1 Change Detection Techniques

In remote sensing, information on LULC change is critical for its practical uses in numerous applications such as forest disturbance (deforestation and degradation), damage assessment, disaster monitoring and management, urban expansion, planning, and land management (Hussain et al., 2013). Remote sensing data has become a major source of data for change detection studies because of its digital format that is suitable for computation, wide range of spectral and spatial resolutions, and its high temporal frequency (Chen et al., 2012). A lot of research, as documented in reviews such as Coppin et al. (2004), Lu et al. (2004), Alqurashi and Kumar (2013), and Hussain et al. (2013), has used remote sensing data to develop change detection techniques. These techniques use multi-temporal remote sensing datasets to undertake a qualitative analysis of the temporal effect of the phenomena, and then quantify the changes (Hussain et al., 2013). The underlying principle of these methods is that they analyse the process and trend of changes by monitoring ground objects based on time series (continuous) remote sensing observation data (Alphan, 2011). Change detection is the process of identifying differences in the state of an object or phenomenon through observing it at different times (Singh, 1989).

Change detection techniques can be classified into three major categories, namely: characteristic analysis of spectral type, vector analysis of spectral changes and time series analysis. Of these, the time series analysis method is relatively the most popular. Broadly, it analyses the process and trend of changes by monitoring ground objects based on time series (continuous) remote sensing observation data. Below is a summary of some major methods relevant to this research. In terms of when change analysis is performed, they can further be broadly classified into either pre or

post-classification comparison methods (Coppin et al., 2004).

Common procedures for the pre-classification approach that are based on image algebra include image differencing (Sohl, 1999; Xu et al., 2009), band rationing (Civco et al., 2002; Alphan, 2011), regression analysis (Hussain et al., 2013), change vector analysis (CVA) (Johnson and Kasischke, 1998), tasselled cap transformation (Rogan et al., 2002), direct multi-date classification (Singh, 1989), vegetation index differencing (Howarth and Wickware, 1981; Podeh et al., 2009) and principle component analysis (PCA) (Hartter et al., 2008). These methods are built around the premise that change in LULC result in differences in the pixel reflectance values between the dates of interest. However, Ridd and Liu (1998) and Dewan and Yamaguchi (2009) noted that one major shortcoming for most of these approaches is that even though they are effective for locating change (simple binary change versus non-change), they are not capable of identifying the nature of change (detailed 'from-to' change).

On the other hand, classification-based comparisons approaches are also as popular, and these include post-classification comparison and multi-date classification. The latter is sometimes referred to as the composite classification. These techniques examine changes over time between independently classified land cover data, and thus give a detailed 'from-to' change information (Jensen, 2000; Coppin et al., 2004; Hussain et al., 2013). Of the two, the post-classification comparison approach is the most popularly used Lu et al. (2004). The 'from-to' change maps are particularly useful for showing the change location, its magnitude and its nature (Howarth and Wickware, 1981). In addition, this technique can be applied even when using data acquired from sensors with different spatial, temporal and spectral resolutions (Dewan and Yamaguchi, 2009; Coppin et al., 2004). However, one of the major disadvantages associated with this approach is that the accuracy of the resultant LULC change maps depends on the accuracy of the individual classification, meaning that such techniques are subject to error propagation (Jensen, 2000; Dewan and Yamaguchi, 2009).

Table 6.1: Summary of the characteristics of commonly used change detection techniques.

Application	Most commonly used techniques	Examples, case studies
1. LULC change	Image differencing, image ratioing, NDVI, post-classification, hybrid change detection, CVA, PCA, ANN, DT, GIS	Sohl (1999), Ridd and Liu (1998), Baronti et al. (1994), Zhou et al. (2009), Jensen et al. (1999), Allen and Kupfer (2000), Miettinen et al. (2011), Foody et al. (1995)
2. Urban change	Image differencing, post-classification, hybrid change detection, PCA, GIS, chi-square, image fusion	Lu et al. (2005), Liu (2004), Zhang et al. (2002), Du et al. (2012)
3. Environmental change	NDVI, ANN, CVA, post-classification, image differencing	Dhakal et al. (2002), Lu et al. (2005), Howarth and Wickware (1981), Liu and Lathrop (2002)
4. Vegetation change	NDVI, CVA, image differencing, post-classification	Lu et al. (2005), Schoppmann and Tyler (1996), Jin and Sader (2005), Xie et al. (2008), Thayn (2012), Healey et al. (2005)
5. Landscape change	Post-classification, GIS	Dhakal et al. (2002), Taylor et al. (2000), Zheng et al. (1997), Kepner and Watts (2000)
6. Deforestation	Post-classification, NDVI, image differencing, PCA	Lu et al. (2005), Lucas et al. (1993), Wulder et al. (2004), Wilson and Sader (2002), Hayes and Sader (2001)
7. Wetland change	Post-classification, GIS	Munyati (2000), Jensen et al. (1993)

The last decade saw a huge hike on the use of machine-learning techniques. These include ANNs, SVMs, and DTs. These three are non-parametric, with ANNs being supervised, while the other two make no assumption on the distribution of the data. They have been shown to be very good for detecting change, but still do have their own limitations which are discussed in detail below. Table 6.1 summarizes the commonly used change detection techniques and examples of applications where they have been used. For the LULC change application, notably, a larger proportion of the techniques are applicable, and have been used to detect change, as documented in the examples of literature given. This research, being one on detecting LULC change therefore reviews a wide range of these techniques.

Pre-classification Approaches

Pre-classification techniques are classed into simple algebra techniques and those called transformation techniques. Those categorised as algebra techniques include image differencing, image ratioing, image regression, vegetation index difference, vegetation index ratioing, and change vector analysis. Under transformation techniques, there is principal component analysis, chi-square transformation, tasselled cap and Gramm-Schmidt (Lu et al., 2004; Hussain et al., 2013). They are amongst the first techniques of change detection, and is relatively simple and easy to implement and interpret (Podeh et al., 2009), providing locational information on areas of spectral change (Sohl, 1999). However, limitations such as their incapability to provide sufficient information on the trend and direction of change (Allen and Kupfer, 2000; Dhakal et al., 2002; Lucas et al., 2010) has made them mostly inappropriate for change detection.

Classification Approaches

This group of techniques are based on classified images, in which good quality classification results are reliant on the quality and quantity of the training sample data (Lu et al., 2004). The four most common classification techniques are unsupervised change detection, post-classification comparison, expectation-maximisation algorithm (EM), and spectral-temporal combined analysis. Of these, only post-classification comparison has been widely used. Unsupervised change detection approach is quick and does not require a-prior information from the user. However, the difficulty associated with identifying and labelling change trajectories seems to make it less attractive as well. Another method that has also been used is the sample-based change reference data method.

However, the method has only been widely used to quantify multi-epoch change (Li et al., 2009; Zhou et al., 2011; Arav and Filin, 2012; Devries et al., 2015; Carroll et al., 2016).

Post-classification Comparison It is a relatively simple change detection analysis technique based on a classification (Kandare, 2011), useful for extracting LULC information. The technique requires the comparison of independently produced classified changes (Alqurashi and Kumar, 2013) at both ends of the time interval of interest (Coppin et al., 2004). Therefore, each image of multi-temporal images is classified separately, pixel by pixel or segment by segment, on two dates. This, according to Alqurashi and Kumar (2013), minimises the atmospheric and sensor differences between the two dates, and further minimise error that may be introduced by poor registration between multi-date images. Thereafter, the resulting independent classifications are compared. Corresponding pixels or objects having the same category label means that pixel/object has not changed, or else the pixel has changed.

Literature is rich with studies that used post-classification comparison (Liu and Lathrop, 2002; Yuan et al., 2005; Shalaby and Tateishi, 2007; Dewan and Yamaguchi, 2009; Afify, 2011; Al-doski et al., 2013; Carreiras et al., 2014), and Alqurashi and Kumar (2013) argues that it is the most commonly used technique in change detection. Advantages of the technique advanced by Mas (1999) include that it is amongst the most accurate procedure and has the ability to indicate the nature of change, thus superior over pre-classification approaches. It does not only ascertain the spatial distribution of the change, but further detail the nature of the change. Kandare (2011) puts it as providing spatial information on the transition from one class to another of the pixel/object. However, Alqurashi and Kumar (2013) is quick to point out that a great deal of time and expertise is required for the performance of post-classification comparison. It has high requirements for a reasonable classification of categories. Furthermore, the review noted that the quality of the classified image for each date affects the final accuracy. It is a method very sensitive to classification errors, and therefore dependent on the accuracy of the classification results (Kandare, 2011).

Machine Learning Techniques

Machine learning is a computational approach to data partitioning and categorization that is based on the idea of learning patterns in datasets (Lippitt et al., 2008). Commonly used techniques in remote sensing based change detection include ANN, SVM, DT, and random forest classifiers,

and these have been applied to both multispectral and hyperspectral data in a wide range of applications (Huang and Jensen, 1997; Foody and Mathur, 2004; Ham et al., 2005; Lippitt et al., 2008; Fauvel and Benediktsson, 2008; Camps-Valls, 2009; Otukei and Blaschke, 2010). Most of these classifiers are explained in Chapter 5, but are also discussed below in their application for change detection.

Decision Trees A DT classifier is technique that performs detection using an algorithm that employs a sequence of decisions to label an unknown pattern (Jensen, 2000). It may be used as a manual or automated technique, with the former noted to be time-consuming and likely to fail in giving satisfactory results for large number of classes and where there is spectral overlap between classes (Alqurashi and Kumar, 2013). Another limitations with this technique is that it does not produce change direction or a change matrix. It is also very sensitive to the training data quality and the number of training samples per class. Therefore, classes can be over-trained, which makes the model not applicable in other environments or even different time periods (Lippitt et al., 2008; Hussain et al., 2013). Other limitations include that it does not search for an optimal fit, and it can grow very big in size, making interpretation hard.

Still, it has its advantages, and these include that it is non-parametric, meaning a user does not have to select a threshold or a training sample. It also does not make any assumptions on the distribution of data. Moreover, this technique is capable of providing the rule-sets for each of the change and non-change classes. Computationally, it tends to be the fastest algorithm under machine learning (Otukei and Blaschke, 2010).

Examples of studies that used DT include Xu et al. (2005) where a regression approach was used to determine class proportions within a pixel so as to produce soft classification from remote sensing data. The study found that the approach had the ability to explore the complex relationships between bands and classes, then identify the most useful combination of bands that increases the class separability between any two classes. Yang et al. (2012) used a 20 years Landsat time series data to detect change in the LULC of Inner Mongolia Autonomic Region of China. Otukei and Blaschke (2010) analysed the potential of DTs as a technique for data mining, using a 1986 and a 2001 Landsat TM and ETM+ datasets, respectively. The study compared DT with SVM and ML, and concluded that the DT performed better than both ML and SVM. Other examples include Friedl and Brodley (1997), Pal and Mather (2003), Pal (2005), Waske and Benediktsson (2007), García et al. (2011) and Vieira et al. (2012).

Artificial Neural Networks ANN belongs to the classification-based change detection category, non-parametric and does not make any prior assumptions about data distribution and dependency (Alqurashi and Kumar, 2013; Hussain et al., 2013). It requires training and testing classifications (Coppin et al., 2004; Lu et al., 2004; Alqurashi and Kumar, 2013; Hussain et al., 2013). The training sample is used to train the algorithm (Lu et al., 2004), which is then used to build relationships (networks) between the input (image) and the output nodes (changes). Thereafter, the trained network is applied to the main dataset (remote sensing and ancillary data) to detect change (Gopal et al., 1996; Dai and Khorram, 1999; **?**). Its main advantage is its ability to provide better change detection results when LULC classes are not normally distributed (Lu et al., 2004). However, its weakness is that it is sensitive to the amount of training data that is used and requires a long training time.

Sadeghi et al. (2013) evaluated the results of modelling data normalisation using ANN in all six bands of Landsat image, a step that produced more optimum results compared to those of normalization with a linear model. ANN could model and decrease the effect of different effective parameters in imaging conditions. A study in the state of Johor in Malaysia used ANN and ML to classify different LULC types and to detect change based on 2000 and 2009 Landsat images, and found ANN to give the highest classification accuracy (Deilmai et al., 2014). Other studies that used ANN for change detection include Foody et al. (1995), Benediktsson and Sveinsson (1997), Jensen et al. (1999), Brown et al. (1999), Liu and Lathrop (2002), Pal and Mather (2003), Qiu and Jensen (2004), and Dixon and Candade (2008).

Support Vector Machines SVM is particularly appealing in remote sensing because it is non-parametric, does not need any prior assumptions on the distribution of the data, and due to its ability to generalize well even with limited training samples, a common limitation for other techniques (Vapnik, 1999; Wijaya and Gloaguen, 2007; Huang et al., 2008; Alqurashi and Kumar, 2013). Their ability to successfully handle small training data sets and further often produce higher classification accuracy than traditional methods (Mantero et al., 2005; Mountrakis et al., 2011) has further made the classifier widely used. A theoretical advantage of SVM over other machine learning algorithms is that it is designed to search for an optimal solution to a classification problem, whereas others algorithms such as DT and ANN are designed to find a solution (which may or may not be optimal) (Huang et al., 2008). This advantage has been demonstrated in a number studies which found SVM to generally produce more accurate results than other machine learning

algorithms. Examples include Burges (1998), Huang (2002), and Nitze et al. (2012).

Other examples of the application of the technique for change detection mapping purposes include Huang et al. (2008), a study that demonstrated that SVM is especially effective for mapping major disturbances with low commission errors. In Algiers, Algeria, Nemmour and Chibani (2006) compared a SVM-based change detection method with different kernel functions to ANN. Having used both statistical and visual evaluations, experimental results showed SVM to have yielded better performance over ANN, both in terms of accuracy and generalization. However, other studies found the other algorithms to outperform SVM. For example, Otukei and Blaschke (2010) concluded that the DT performed better.

K-NN Classifier For the K-NN classifier, given an unknown feature vector x and a distance measure, out of a number of N training vectors, it seeks to identify the k nearest neighbours, irrespective of the class label. Theodoridis and Koutroumbas (2003) states that for a two class problem, the k is chosen to be odd, and not to be a multiple of M classes. Then out of these k samples, it identifies the number of vectors (k_i) that belongs to the class ω_i, $i = 1, 2, \ldots, M$. Common measures include the Euclidean and Mahalanobis distance. The simplest version of the algorithm is to assign the each feature vector to the class of its nearest neighbour, and the rule has been reported to perform well with a large training sample.

Naidoo et al. (2012) cited a major drawback associated with the K-NN techniques as being the complexity of searching for the optimal nearest neighbour(s) among the N training samples, a problem that becomes particularly severe in high-dimensional feature spaces (Theodoridis and Koutroumbas, 2003).

Random Forests Algorithm The emergence of the random forests approach was seen as an improvement as concepts such as multiple DTs, bootstrap aggregation (bagging) and internal cross-validation were introduced, which led to improved results, ease of use, and overcoming of the issue of over-fitting (Ismail et al., 2010; Naidoo et al., 2012).

The random forests classifier is a relatively new entry to the remote sensing field of change detection and according to Breiman (2001), was specifically designed to produce accurate predictions that do not over-fit the data. However, it has become popular with many scientific fields because it can cope with with both 'small or large' problems, complex interactions, as well as even highly

correlated predictor variables (Strobl et al., 2008). It consist of a combination of tree classifiers where each classifier is generated using a random vector sampled independently from the input vector. The number of features used at each node to generate a tree and the number of trees to be grown are two user-defined parameters required to generate a random forests classifier. Then each tree casts a unit vote for the most popular class to classify an input vector (Breiman, 1999). Random forests draws bootstrap samples to construct multiple trees. Each tree is grown with a randomized subset of predictors, which gives the technique its name 'random' forests. A large number of trees are grown, hence a 'forest' of trees. The number of predictors used to find the best split at each node is a randomly chosen subset of the total number of predictors. The trees are grown to maximum size without pruning, and aggregation is by averaging the trees (Prasad et al., 2006). At each node, only selected features are searched for the best split. Thus, the random forests classifier consists of N trees, where N is the number of trees to be grown, which can be any value defined by the user.

To classify a new dataset, each case of the datasets is passed down to each of the N trees. The forest chooses a class having the most out of N votes, for that case (Pal, 2005). Out-of-bag (OOB) samples can be used to calculate an unbiased error rate and variable importance, which eliminates the need for a test set or cross-validation (Pal and Mather, 2003; Chan and Paelinckx, 2008; Ismail and Mutanga, 2011). There are many approaches to the selection of attributes used for DT induction and most approaches assign a quality measure directly to the attribute. The classifier uses the Gini Index (Breiman, 1999) as an attribute selection measure, which measures the impurity of an attribute with respect to the classes. For a given training set T, selecting one case (pixel) at random and saying that it belongs to some class Ci, the Gini Index is given as:

$$\sum \sum_{j=-i} (f(C_i, T)/|T|)(f(C_j, T)/|T|) \tag{6.1}$$

where $f(C_i, T)/|T|$ is the probability that the selected case belongs to class C_i.

A very useful feature of random forests that sets it apart from the rest of machine learning approaches is that it does not suffer from over-fitting. This is because a large number of are grown which leads to very limited generalisation error (true error of the population as opposed to the training error only) (Adjorlolo et al., 2012). Unlike other machine learning algorithms, random forests algorithm is relatively easy to implement, requiring only two parameters. These are the

number of trees to grow and the number of variables to split at each node. Another important feature of random forests is that it offers variable importance measures embedded in its computation and it can perform both binary and multi-class classification tasks. The variable importance measures provide the researcher with valuable insight to explore the effect of any predictor variable on the response variable (Strobl et al., 2008; Adjorlolo et al., 2012).

Pal (2005) found random forests to produce results comparable to SVM, but further observed that the classifier was able to handle categorical, unbalanced data, and data with missing values, an attribute SVM still lacked. A predictive vegetation mapping study by Prasad et al. (2006) concluded that random forests outperformed Regression Tree Analysis (RTA), Bagging Trees (BT) and Multivariate Adaptive Regression Splines (MARS). Generally, random forests seem to be a popular classifier for the SADC region as a number of studies preferred it, and generally reported good results. Examples include Abdel-Rahman et al. (2009), Ismail and Mutanga (2011), Mutanga et al. (2012), Naidoo et al. (2012), and Adam et al. (2012). Ismail and Mutanga (2011) explored how well random forests could discriminate between healthy trees and those in their early stages of Sirex *noctilio* infestation. For three variables selecting methods, random forests was found to produce better classification results for all three compared to competing boosting trees. Chan and Paelinckx (2008) assessed random forests against another tree-based ensemble classification algorithm (Adaboost), based on standard classification accuracy, training time and classification stability. The study found that these two algorithms gave almost the same overall accuracy (approximately 70%), and outperformed ANN (63.7%). Moreover, it was found that random forests was faster in training and more stable than the other two algorithms.

For vegetation and LULC mapping, random forests has been used by studies such as Akar and Güngör (2012) Naidoo et al. (2012), Gong et al. (2013), Hayes et al. (2014), Mishra and Crews (2014), Eisavi et al. (2015), and Kulkarni and Lowe (2016). Most of these LULC studies found the approach to be very effective, yielding quality results. For example, Hayes et al. (2014) classified riparian areas within a 225 km^2 study area, using random forests, and produced a land cover map with overall accuracy of 81% and a kappa coefficient of 79%. Eisavi et al. (2015) used random forests to classify for LULC from Landsat 8 imagery, for four scenarios, and the products had overall accuracies of 86.48%, 82.26%, 90.63%, and 91.82% respectively. Akar and Güngör (2012) evaluated the performance of random forests by comparing its classification to results obtained from Gentle AdaBoost (GAB), SVM and MLC algorithms. Preliminary results indicated that random forests gave higher classification accuracies than the other methods for Quickbird and

Ikonos images.

Hybrid Change Detection Approaches

A hybrid technique uses two or more techniques to detect a change. Two types of hybrid techniques exist, and these are procedure-based hybrid analysis and result-based hybrid analysis (Hussain et al., 2013). The former involves the use of different detection techniques in different detection procedures, while the latter involves the use of different change detection techniques successively and then analysing the results. This technique makes full use of the benefits of many techniques to obtain significant change detection results (Min et al., 2006). On the other hand, the hybrid change detection technique is complex, difficult to perform and not very effective because it depends on the characteristics of other techniques (Lei et al., 1998; Tademir et al., 2012; Wondrade et al., 2014).

According to Alqurashi and Kumar (2013), combining techniques has been used in some studies to increase classification accuracy as well as to avoid the leak of training sample sets for image classification. Li and Yeh (1998) used a hybrid of supervised maximum likelihood classification with combined principle component analysis to prevent overestimation of change. There are other numerous literature on how the technique has been used, and some studies reported an improvement that individual techniques were not able to produce in singularity. For example, in a study in Slovenia, Poland and Ukraine that combined supervised and unsupervised classification to derive a land cover map from three Landsat Thematic Mapper Plus (ETM+) images gave more accurate results than when using individual techniques (Kuemmerle et al., 2006). Zhang et al. (2007) concurred with Kuemmerle et al. (2006) as the study reached a similar conclusion when using hybrid change detection and a DT classifier based on a data mining algorithm, as better accuracy was achieved compared to when individual change detection techniques were used.

Wondrade et al. (2014) used a hybrid approach in order to reduce spectral confusion due to high variability of land cover. In an almost identical study, McDermid et al. (2008) combined pixel-based and object-based techniques to reduce noise in change detection, as well as the small and spurious changes introduced by the inconsistent delineation of objects. In 11 municipalities in the northern and western parts of the Former Yugoslav Republic of Macedonia (FYROM), Al-Khudhairy et al. (2005) applied pixel-based principle component analysis and image differencing to VHR Ikonos imagery, and the resultant change images analysed using object-based image analysis,

a step reported to have improved the pixel-based change detection. Petit and Lambin (2001) used this method to increase comparability between land cover maps coming from panchromatic aerial photographs and multi-spectral SPOT XS. Yu et al. (2010) performed an object based classification using a SVM and compared the objects with land use GIS data. Walter (2004) integrated GIS and maximum likelihood classification, and concluded that results of traditional change analysis can be effectively interpreted by supplementing them with object-based post-classification. Other examples of the integration of different studies that used the hybrid approach include Niemeyer and Nussbaum (2006), Niemeyer and Nussbaum (2007), Hofmann et al. (2008), and Gamanya et al. (2009), and most of these studies highlight the that hybrid techniques avail, and how one can adopt and adapt a combination of any techniques as per need.

6.2.2 Evaluation of Change Detection Techniques

A good change detection method is one that will give accurate and detailed change results, accounting for the magnitude and direction of change (Alqurashi and Kumar, 2013). Table 6.2 summarises the advantages and disadvantages of the commonly used techniques, highlighting their capabilities and their limitations. Another aspect that is included is the hardness associated with implementing them. One major limitation for most of the older algebra-based approaches is that these are not capable of providing sufficient information on the trend and direction of change (Allen and Kupfer, 2000; Dhakal et al., 2002; Lucas et al., 2010). However, some studies found them to be superior over some classification methods and machine learning methods for detecting change (Liu, 2004; Berberoglu and Akin, 2009). But the lack of change trends and direction is a limitation for monitoring purposes, which makes this group of techniques inappropriate, especially if used in singularity. On the other hand, numerous techniques such as post-classification, change vector analysis, ANNs, SVMs and random forests are capable of providing more information on the direction and trends of change (Kandare, 2011; Deilmai et al., 2014). As already highlighted, this is information critical both for initial change detection and for long-term monitoring purposes.

Another limitation with techniques with most of the algebra techniques is that they require a threshold that is used to select change pixels or features from non-change (Lu et al., 2002; Pu et al., 2008; Nandy et al., 2011). It is therefore, critical to select an appropriate threshold as it determines how good the change results will be. However, the process of establishing the threshold is quite difficult and can be time-consuming, as noted by Coppin et al. (2004) and Alqurashi and

Kumar (2013). This is where a good technique like change vector analysis which is otherwise capable of giving change results that gives the direction and trends (Lu et al., 2004), tend to fall short.

Techniques that are capable of providing information on change direction and change, yet not limited by the need to select a threshold include both classification and machine learning techniques, and these are post-classification, ANN and SVM. The selection of the most appropriate technique factors in both its weaknesses and strengths. For example, of these four, post-classification is known to minimise the effects of the atmosphere, the sensor used to acquire data, and environmental differences between images resulting from different times of acquisitions (Singh, 1989). However, the technique is greatly affected by errors, which then usually requires a lot of time and effort dedicated to reducing errors. Unfortunately, this cannot be guaranteed all the time. According to Alqurashi and Kumar (2013), the confusion matrix is commonly used to account for omissions and commissions. Since it uses two independent classifications from two independent images acquired at different times and under different conditions, each of these tend to contribute its own independent errors into the derived change detection maps (Singh, 1989).

An advantage of the transformation category is in their ability to reduce data redundancy between bands and emphasizing different information in derived components (Singh and Harrison, 1985; Liu and Lathrop, 2002). In the transformation category, principle component analysis is most often used for detecting change/non-change information (Baronti et al., 1994; Tso and Mather, 2009). The Chi-square method is relatively less frequently used in practice due to its relative complexity compared to principle component analysis (Howell, 2009). Also, Chi-square is not readily available in most of commercial remote sensing image processing software. But overall, like the algebra techniques, these transformation techniques do not provide detailed change matrices, and further require the user to select thresholds to identify change (Ridd and Liu, 1998). Another disadvantage is the difficulty in interpreting and labelling the change information on the resultant transformed images.

The classification category methods are based on the classified images, in which the quality and quantity of training sample data are crucial to produce good quality classification results (Lu et al., 2004). Post-classification comparison is the most commonly used of these approaches for change detection, but the difficulty in classifying historical image data often seriously affects change detection results derived from it (Kandare, 2011). The major advantage of these methods is their

capability to provide a matrix of change information and reducing external impact from atmospheric and environmental differences between the multi-temporal images (Alqurashi and Kumar, 2013). However, Kandare (2011) noted that selecting high-quality and sufficiently numerous training sample sets for image classification is often difficult, in particular for historical image data classification. The time-consuming and difficult task of producing highly accurate classifications often leads to unsatisfactory change detection results, especially when high quality training sample data are not available (Lu et al., 2004), which was the case for Liwale.

Machine learning techniques, being non-parametric, have had the edge over algebra and transformation and classification techniques. The main issues with ANN, however, include the 'hidden' layer not properly known, which has earned it the nickname of a black-box (Qiu and Jensen, 2004; Xie et al., 2008; Liu et al., 2012). In addition, a good change result is very reliant on the training data provided. DT suffers from high dependence on the training data as well such that, as Lippitt et al. (2008) states, they can be over-trained so that the model is only specific for that area it has been developed for, and thus not transferable to other environments. This work sought to develop a model that can be taken across to other savanna environments, both within Tanzania and other SADC regions, as well as other forest types within the country. This, therefore, made DT technique not the preferred approach. Another challenge associated with DT approach is that it can grow very large in size, making its interpretation very difficult (Hussain et al., 2013).

Like ANN, SVM requires a training and a testing sample. The provided training sample is the determinant of how good the change results be; it is quite sensitive to the amount of provided training sample (Huang et al., 2008; Mountrakis et al., 2011). This could lead to either over-fitting or under-fitting if the sample is inappropriate (Alqurashi and Kumar, 2013). They also tend to require a long time for training while overall, these techniques were not common in remote sensing software (Lu et al., 2004). On the contrary, random forests technique is not affected by the sample size; it is not affected by the over-fitting problem (Adjorlolo et al., 2012). This set the technique as superior and the most appropriate for change detection in these highly heterogeneous savanna forests of Liwale. It was chosen on account of its superior generalization capacity, insensitivity to limited number of empirical training samples, and ability to reduce classification error on a wide array of data structures. Furthermore, it can handle categorical data, unbalanced data, as well as data with missing values, which is still not possible with SVM (Pal, 2005). In addition, random forests is relatively the easiest to implement within the set of machine learning algorithms, requiring only two parameters.

The hybrid change detection method combines the advantages of the threshold and classification methods (Li and Yeh, 1998; Wondrade et al., 2014). Threshold methods are often used to detect the changed areas, then classification methods are used to classify and analyse detected change areas using the threshold method (Lu et al., 2004). For any method using T_1 and T_2 land cover maps to determine change, the accuracy of those methods can not be better than that of each of the input maps, and is often lower. An approach based on image-segmentation and rule-based classification is potentially such an improved methodology, perhaps in an integrated approach (Civco et al., 2002). This study utilised an almost similar approach as the data was first thresholded for potential change features before these features were classified for change using the selected classifier (random forests). Table 6.2 is a summary of the characteristics of the most commonly used change detection techniques in remote sensing.

Table 6.2: Summary of the characteristics of commonly used change detection techniques.

Technique	Pre vs post	Data used	Disturbance type	Change duration	Selecting threshold	Selecting training sample	Provides change direction	Provides change matrix	Difficulty of implementing
1. Univariate image differencing	pre	satellite imagery	deforestation and degradation	abrupt and progressive	Y	N	N	N	1
2. Image ratioing	pre	satellite imagery	deforestation	abrupt	Y	N	N	N	1
3. Change vector analysis	pre	satellite imagery	deforestation and degradation	abrupt and progressive	Y	N	Y	N	3
4. Post-classification comparison	post	satellite imagery	deforestation and degradation	abrupt	N	Y	Y	Y	2
5. Principle component analysis	pre	satellite imagery	deforestation and degradation	abrupt	Y	N	N	N	2
6. Artificial neural networks	pre	satellite imagery and ancillary data	deforestation and degradation	abrupt	N	Y	Y	Y	5
7. Support vector machines	pre	satellite imagery	deforestation and degradation	abrupt	N	Y	Y	Y	5
8. Image regression	pre	satellite imagery	deforestation and degradation	abrupt	Y	N	N	N	1
9. Decision tree	pre	satellite imagery	deforestation and degradation	abrupt and progressive	N	N	N	N	3
10. Image fusion	pre	satellite imagery	deforestation	abrupt	Y	N	N	N	2
11. Hybrid change detection	pre	satellite imagery	deforestation	abrupt or progressive	Y	N	N	N	3
12. Random forests	semi	satellite imagery	deforestation and degradation	abrupt and progressive	N	Y	Y	Y	2
13. Chi-square	pre	satellite imagery	deforestation	abrupt	Y	N	N	N	3
14. GIS	pre	satellite imagery and ancillary data	deforestation and degradation	abrupt	N	N	Y	Y	4

6.3 Methods

The approach employed in this study was a classification-to-image approach (Figure 6.1). It first uses a single date (wet season) image to establish a forest/non-forest baseline (classification) for the area of interest, and the resultant classification together with an image of a different date is used to detect change. Unlike the post-classification approach that classifies both images independently and then compares the two independent classifications, this approach first establishes the best possible forest baseline of an area, then seeks to find any change on the baseline using the image of a later (different) date.

The advantage with this approach is that, unlike the post-classification approach where errors from both classifications gets propagated into the change detection process and product, only errors from the baseline product get propagated into the change map. Moreover, due to the difficulties associated with acquiring a full data coverage of an area of interest whenever needed due to acquisition schedule of the sensor of interest (in this case being RapidEye), cloud cover and seasonality, classification-to-classification comparison eliminates the flexibility of being able to easily update the baseline just over the available scene(s) at any given point. The classification-to-image approach means that immediately data becomes available over any part of a site of interest, no matter how small or big the coverage, the existing baseline can be immediately updated over those portions or areas that has received new data coverage.

Satellite images of Liwale acquired in May 2012 and February 2014 were well suited for comparison (change detection), and Liwale is relatively of a flat terrain. Both the images were carefully selected to ensure that they were both acquired during the wet season (summer) period, with minimal phenological difference and seasonal factors. The approach safeguards against limitation arising from image differences as the forest baseline is classified for independently of the image it is compared against during change detection. Therefore, during the change detection process, the spectral range of values for the image being used is not affected by the baseline.

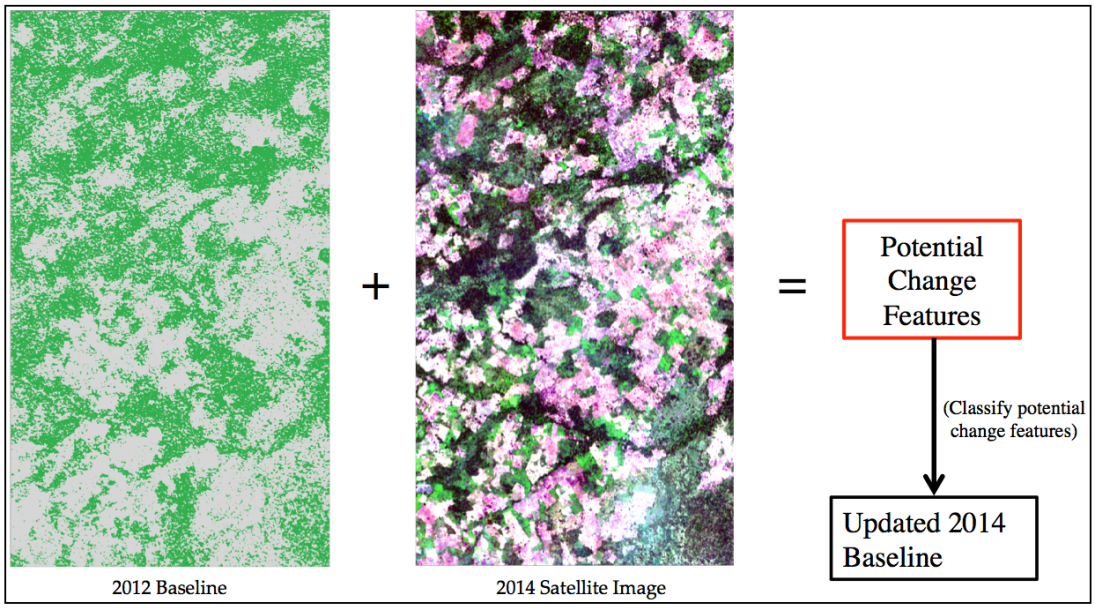


Figure 6.1: An illustrative diagram of the flow of steps followed to identify potential change features, using RapidEye data.

Using the 2012 forest baseline, together with the new image (RapidEye 2014 wet season image), potential change features were identified through thresholding of spectral bands and indices identified as best for separation of these features. Thereafter, these features were classified to produce a change map and an updated 2014 forest baseline of the area using random forests classifier.

6.3.1 Indices and Data Thresholding

Data thresholding was used to identify the potential change features, through the use of Skewness and Kurtosis, which is detailed below. As the baseline was for the year 2012 and the new acquisition imagery was for the year 2014 (only two years apart), only forest to non-forest changes were possible, and thus considered. Forests will not have regrown in the 2-year time period, especially savanna woodlands that have a relatively low tree growth rate as noted by Rossatto et al. (2009) and Tomlinson et al. (2014). For this reason, change from non-forest to forest was not considered.

The thresholding approach was based on an assumption that the data was normally distributed, and the tail of the normal distribution on the post change data are the potential change features. It was hypothesized that a histogram of data for a spectral band or index for a land use/cover class for time 1 (pre change) would be normally distributed, but would no longer be normally

distributed for time 2 (post change) due to some objects having been changed or modified. As it has been shown by studies, change areas/objects have different spectral signatures from those of non-change. The assumption was that the resulting tail on the histogram of the post change data represented the change features/objects within the scene.

To test the hypothesis for the RapidEye data, first, histograms for the 2012 forest class were generated for different spectral bands and indices. Again, using the same bands and indices that had been used on the 2012 imagery, histograms were generated for the 2014 imagery, using only data falling within the 2012 forest spatial extent. Figure 6.2 is an example of the forest class in 2012 and post change in 2014 (with potential change) for CCCI index. The histograms demonstrated that for 2012, the spectral response for the forest class was normally distributed, while for 2014 there was a clear tail, which were thresholded as potential change features.

Having confirmed that the assumption was true that the data was normally distributed for the forest class, a sample of areas that had remained forest (unchanged) in both 2012 and 2014 were manually selected by digitizing polygons in ArcMap software. The same was done for areas of change; that had changed from forest in 2012 into non-forest in 2014. Then for the different indices and image bands that were found to give a good discrimination between forest and change features, histograms were plotted (Figure 6.3). These plots highlighted the possibility of discriminating between forest and change areas that had gone to non-forest, or rather those that had experienced disturbance. Moreover, the plots showed that the tail on the forest histograms from the 2014 statistics was the potential change features.

Figure 6.3 are histograms for CCCI, CIRedEdge, LCI, NDVI, NIR and PSRI. It shows that these indices and image bands had different abilities to separate potential change features from the forest class. Some had a clear separation between the histograms, whereas some had bigger tails, showing a confusion between the forest and non-forest classes. Having proven the normal distribution within data for a class, and that change features differed from the class they were before change, the next step was to threshold the forest class dataset so as to separate the potential change features on the later image.

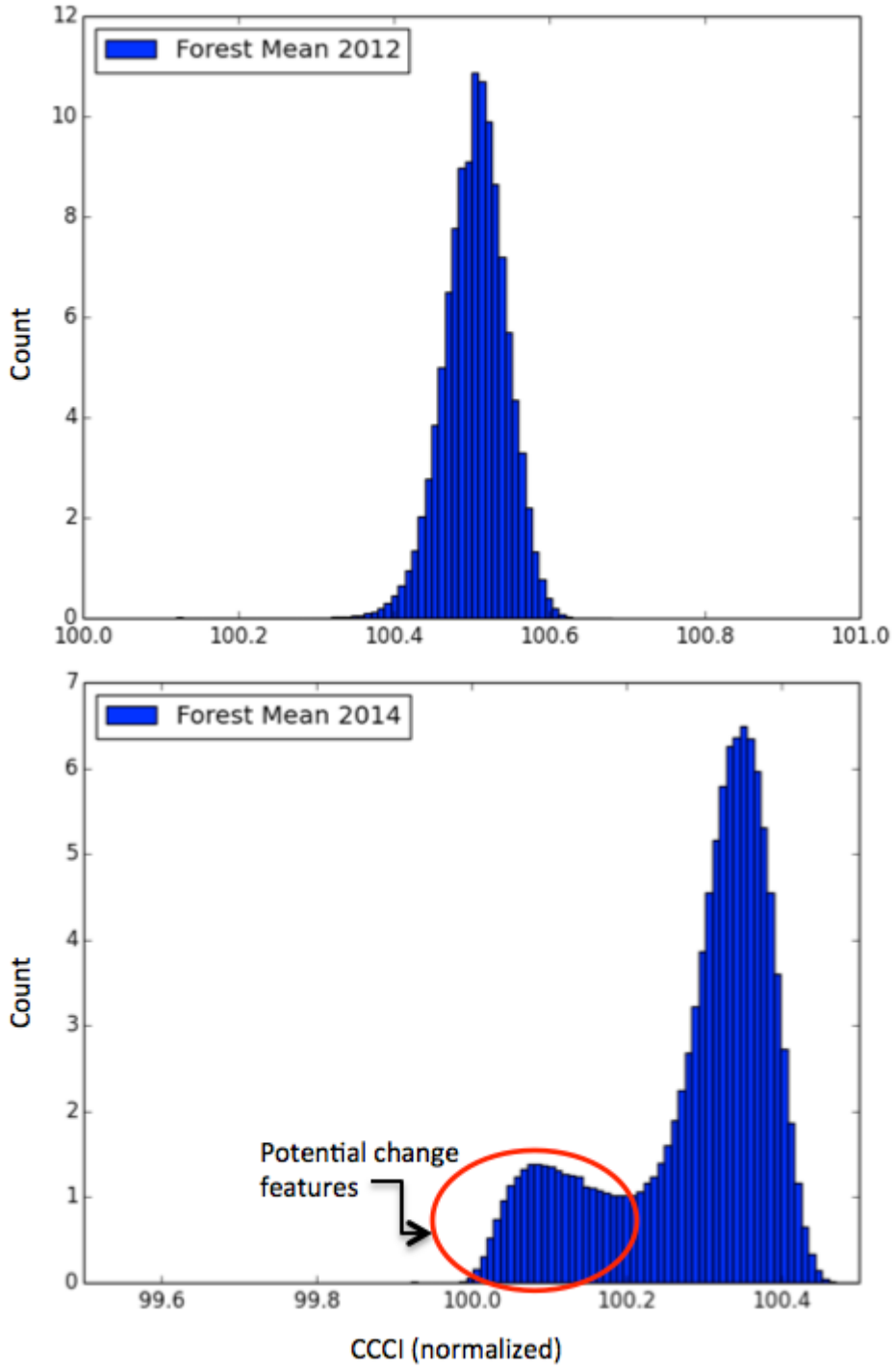


Figure 6.2: Histograms of the forest class in 2012 and post change in 2014, with potential change.

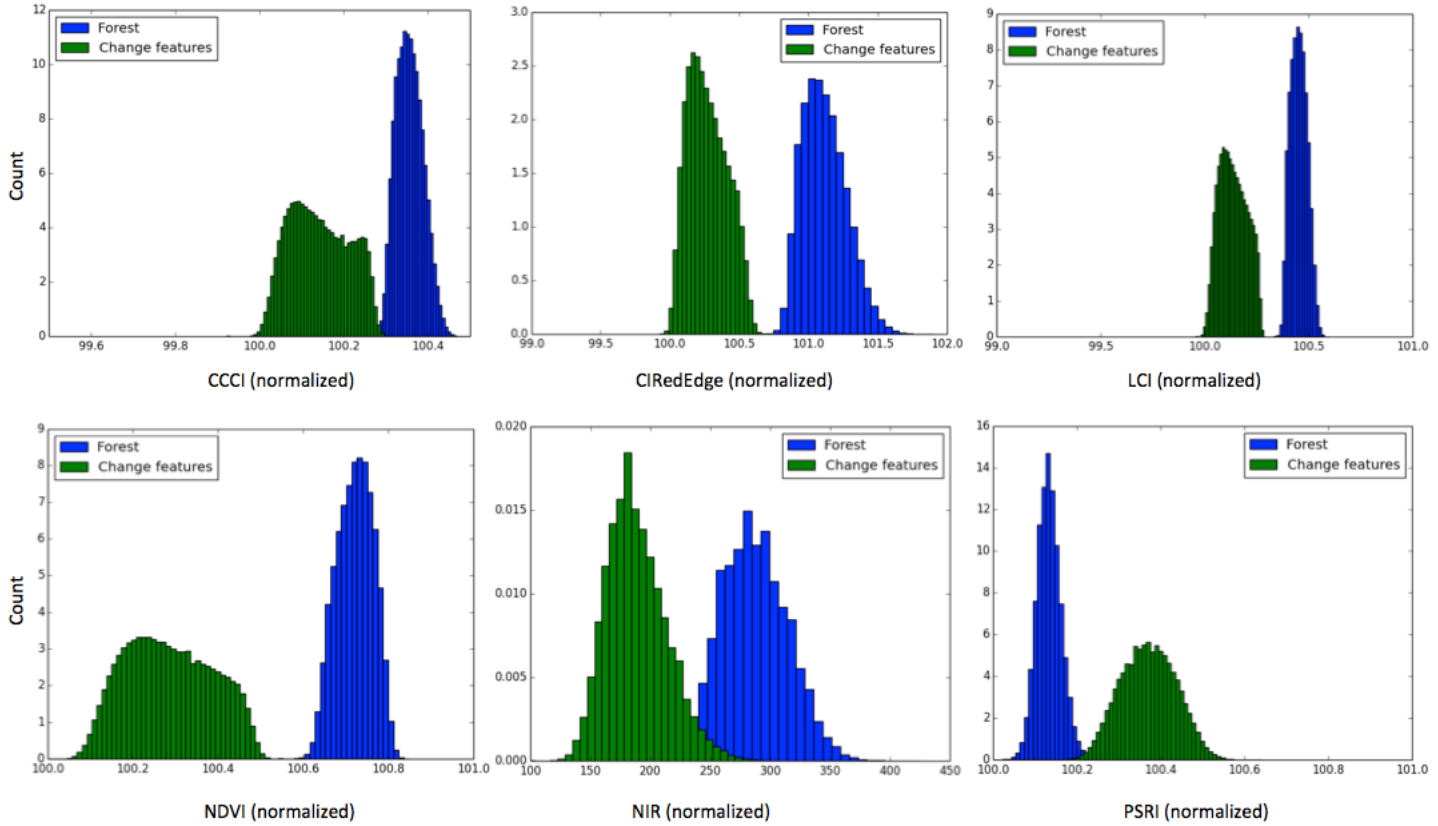


Figure 6.3: Histograms showing the normal distribution and discrimination between the forest class and the disturbed (potential change) features for some indices and spectral bands.

6.3.2 Thresholding Data for Potential Change Features

A fundamental task in many statistical analyses is to characterize the location and variability of a dataset. For normal distributions the metrics skewness and kurtosis can be used to express the normality of the distributions. To threshold the data and separate the potential change features, two thresholding techniques were considered. These were Otsu's thresholding method and Skewness and Kurtosis optimiser.

Otsu's Thresholding Method

Otsu's thresholding is a non-parametric and unsupervised method of automatic threshold selection for image segmentation. It automatically performs a clustering-based thresholding. In an ideal case, the histogram has a deep and sharp valley between two peaks representing foreground and background, respectively, so that the threshold can be chosen at the bottom of this valley (Liu et al., 1991). This algorithm uses a simple approach that assumes the histogram being thresholded is bi-modal. Therefore, an optimal threshold is selected by the discriminant criterion through choosing a threshold that minimizes the intra-class variance between these two sets, a step that was observed by Otsu (1979) to be the same as maximizing the between-class variance. It is a relatively simple procedure, utilizing only the zeroth- and the first-order cumulative moments of the gray-level histogram. For the change scenario in Liwale where there was clearly a tail which was dominated by change features. Otsu's method was therefore considered for thresholding the histogram.

Examples of change detection studies where Otsu's method has been used to threshold the data include Renza et al. (2013), which compared Erreur Relative Globale Adimensionnelle de Synthese (ERGAS) index and change vector analysis methods for detecting change. For the purpose of thresholding the data for change detection, this study compared Huang, Entropy-kapur, Moments, Otsu, Renyi and Shanbhag methods, and concluded that Otsu and Moments were superior for ERGAS index, while for change vector analysis technique, Otsu was the best method for thresholding the data. On the contrary, other studies such as Zhang et al. (2010a) concluded that the method was not complete for the extraction, there were problems of missing detection. Other numerous studies such as Liu et al. (1991), Xu et al. (2014) and Huang et al. (2015), successfully used the method for image segmentation purposes.

Skewness and Kurtosis

Skewness and kurtosis is a descriptor of the shape of a probability distribution. A histogram of a dataset of interest is considered an effective graphical technique of showing the skewness and kurtosis of that dataset. Skewness is a measure of symmetry within the dataset. More precisely, it is the measure of the asymmetry (the lack of symmetry) of the probability distribution of a real-valued random variable about its mean (Pyzdek, 2003). A distribution of a dataset is therefore, symmetric if it looks the same to both sides of the centre point, otherwise asymmetric (Figure 6.4). Normally distributed data has a skewness value of zero, meaning that in real life scenario, normally distributed data should have a skewness value near zero. A negative skewness value means that the data are skewed left, meaning that the left tail is relatively longer than the right tail of the distribution. A positive skewness value indicates that the data are skewed to the right, with a right tail longer than the left tail of the distribution (Groeneveld and Meeden, 1984).

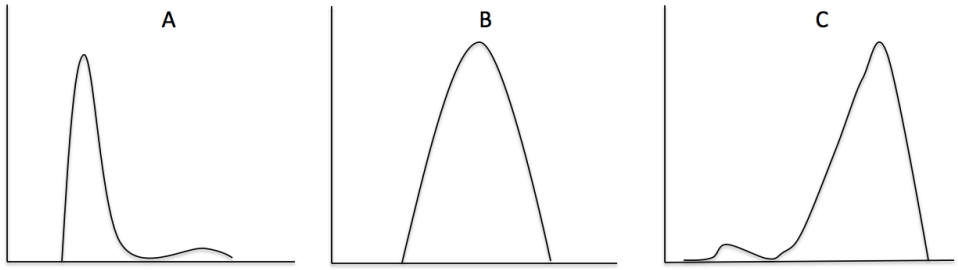


Figure 6.4: A) is an illustration of a histogram with negative skewness, B) is a histogram with a zero value skewness (normally distributed), and C) is a histogram with positive skewness.

Kurtosis on the other hand, measures the heaviness of the tails of a distribution, which Pyzdek (2003) defines as the measure of flatness of the distribution. It is a measure of whether the data are peaked or flat, relative to a normal distribution. A dataset with high kurtosis tend to have a histogram characterised by a distinct peak near the mean, with a sharp decline, and heavy tails, yet one with a low kurtosis tend to have a flat top near the mean rather than a sharp peak (Figure 6.5). In real life scenarios, a uniform distribution is not common, rather an extreme case (Balanda and Macgillivray, 1988).

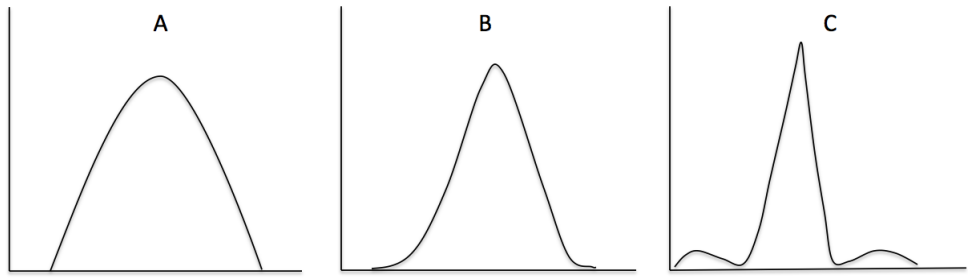


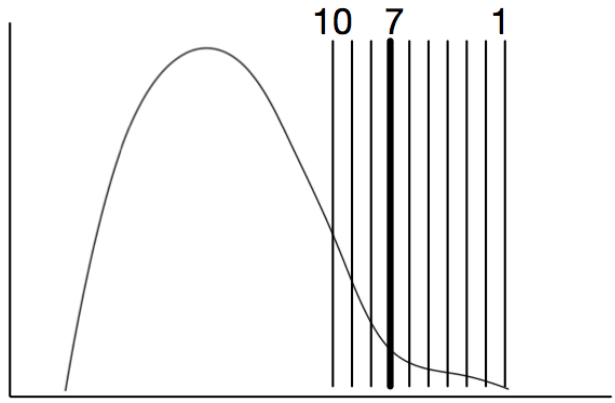
Figure 6.5: A) is an illustration of a histogram with low kurtosis, B) is a histogram with medium kurtosis, and C) is a histogram with high kurtosis.

Examples of a studies that used skewness and kurtosis to threshold data include Mancino et al. (2014), that used two Landsat images two years apart in Basilicata region, southern Italy, to produce their NDVI products. Then the two NDVI images were used to derive an NDVI difference product, from which change was detected through a threshold derived using statistical variables, including skewness and kurtosis. The study used skewness and kurtosis, together with other statistical variables such as the mean, mode, median, standard deviation to test and determine the goodness-of-fit to a normal distribution of the NDVI difference. Chiang et al. (2001) built on the fact that both skewness and kurtosis are susceptible to outliers, and therefore used them as a base to design a projection index for target detection. The results showed that the proposed projection pursuit approach based on the method provided an effective means for target detection. Alvarez-Borrego and Martin-Atienza (2013) used skewness and kurtosis properties of remote sensing images to derive statistical properties of sea surface slopes, and concluded that it was possible to observe sea surface slope through observing skewness and kurtosis of an optical remote sensing image.

Within this study, to identify a threshold using the skewness and kurtosis, an iterative approach was adopted, where the side (i.e., left or right / lower or upper) of the histogram where the change was expected was considered. Each bin that was identified as potential threshold was tested by calculating the sum of the kurtosis and absolute skewness:

$x = |\text{Skewness}| + \text{Kurtosis}$

The lowest value of x therefore identified the optimal threshold to cut the histogram identifying potential change features. Figure 6.6 illustrates this process where for each line (i.e., potential threshold) the skewness and kurtosis is calculated and the minimum is selected (i.e., line 7), and this threshold is used to split the histogram, identifying the potential change features.



Line	Skew+Kurt
1	3.4
2	3.2
3	2.95
4	2.8
5	2.6
6	1.95
7	1.5
8	1.8
9	1.65
10	1.7

Figure 6.6: Demonstrates the iterative thresholds being tested where the minimum is selected as the 'optimal' threshold.

Table 6.3 are the results of each of the two methods compared to a threshold manually computed from the sample data. Skewness and kurtosis was found to be the most consistent with the manually computed threshold values, compared to Otsu's method. Therefore, skewness and kurtosis is the method that was used to threshold the bands and indices that were used to determine the change features as well as potential change features.

Table 6.3: The threshold values from the training set, skewness and kurtosis optimiser, and Otsu's thresholding method.

Index	**Training sample**	**Skewness and Kurtosis**	**Otsu algorithm**
CCCI	0.25	0.26	0.26
NDVI	0.57	0.59	0.44
LCI	0.35	0.35	0.32
CIRedEdge	0.53	0.52	0.56
NIR	240	243	248
PSRI	0.21	0.21	0.16

Having found skewness and kurtosis to be superior over Otsu's thresholding for Liwale environment, the study followed the steps implemented by Mancino et al. (2014), where a histogram for an index or spectral band of the class of interest was thresholded using skewness and kurtosis. The threshold would identify either two or three ranges in the normal distribution: (a) the left tail; (b) the right tail; and (c) the central region of the normal distribution. Features within the tail(s) of the distribution are what were assumed to be characterised by significant vegetation changes, being

termed potential change features. Then the features within the main normal distribution were the regions representing non-change. Figure 6.7 is an illustration of the use of skewness and kurtosis to threshold for potential change features using CCCI. For this index, change was bimodal, data only tailing to the left. A threshold line was fitted using skewness and kurtosis to the graph. A flowchart summarising the steps that were undertaken to perform the identification of potential change features using the skewness and kurtosis optimisation is given in Figure 6.8.

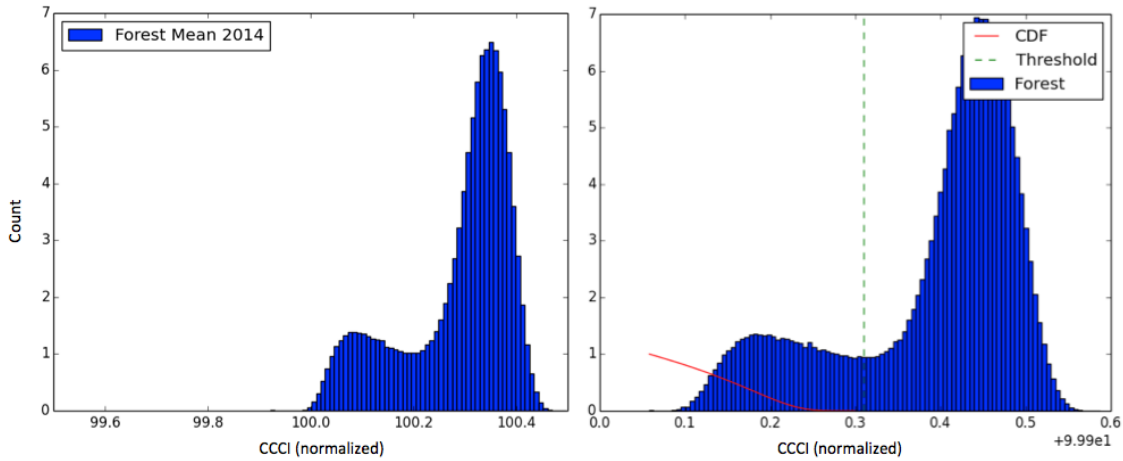


Figure 6.7: Thresholding of potential change features from forest using skewness and kurtosis for CCCI.

To note is that no actual tests were done to establish the amount of outliers. However, since the study used the statistics as an optimisation, then any noise which would cause a skew/outliers in the class reflectance would be identified as a possible change feature. As these features are not directly used as change, but inputted into the classification step, then even though it was possible that the number of potential change features was overestimated, the classification step would likely correct for these if regions of forest with a similar spectral response were present in the scene.

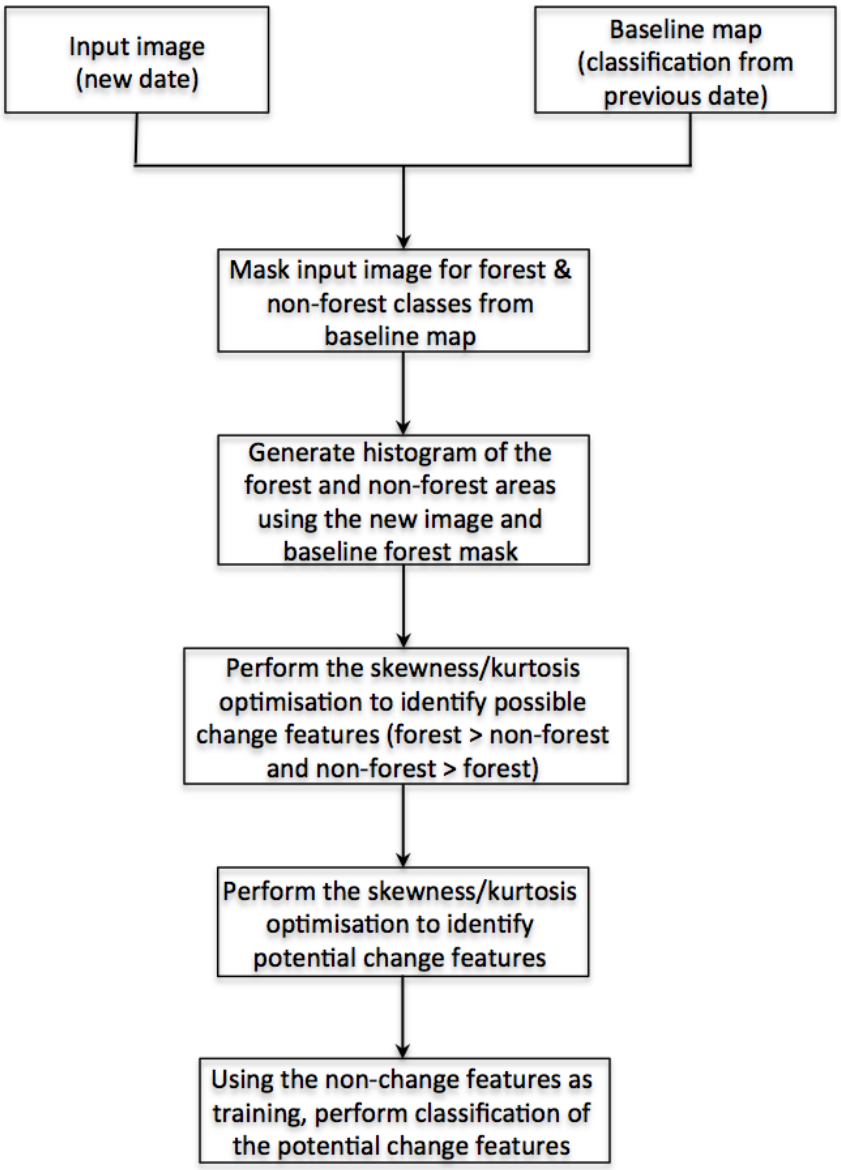


Figure 6.8: A detailed flowchart of the steps undertaken to threshold the data, and identify potential change features, using skewness and kurtosis optimisation.

6.3.3 Classifying Potential Change Features

Once the potential change features were identified, they needed to be classified. The majority of change features were expected to be non-forest, and the remaining non-change features (i.e., existing non-forest extent and forest areas which are not change) can automatically provide the training samples needed to classify the potential change features. However, although the forest

class represents a single discrete part of the feature space, the non-forest class does not as it is made up of many different land covers. Therefore, the classifier used need to be able to handle classes that do not form discrete regions of the feature space.

There are three classifiers that provide this attribute, and these are; random forest, extreme random forest and K-NN. K-NN only considers the K nearest training samples to the unknown sample, and therefore the distribution of the training samples is not important. For random forest and extreme random forest, many independent decision tree classifiers are formed where only a small number of features (e.g., indices) and a sub-sample of the training data are used to build each decision tree, allowing decision trees to be formed in different parts of the feature space for the same class (Naidoo et al., 2012). The three classifiers were tested to determine which was optimum for the savannas forests of Liwale.

Testing for the Optimum Classifiers

The key parameter of the K-NN classifier is the number of neighbours (Theodoridis and Koutroumbas, 2003) while for random forests and extreme random forests, it is the number of trees (Prasad et al., 2006). To test for their performance on the savanna environments of Liwale, a sample comprising of 33381 forest and 39437 non-forest objects were classified for using these three classifiers, for K-NN using different values of k, and different number of trees for random forests and extreme random forests. Table 6.4 are the results of the performance test undertaken on the three classifiers. Table 6.4c shows that the K-NN classifier lack both consistently and accuracy across a range of 2 and 30 k values it was tested for. An increase in k was found to lead to a bigger error for the sample dataset.

Table 6.4a and Table 6.4b show that the random forests and extreme random forests classifiers were more accurate and highly consistent in classifying the testing sample across a range of a number of trees used. Even though extreme random forests was very consistent, on average, about 20 of the 33381 forest sample were being incorrectly classified into non-forest (Table 6.4b). On the other hand, the random forests classifier was both highly consistent and accurate, either correctly classifying the whole sample or misclassifying only one sample object for the whole range of number of trees tested (Table 6.4a). Therefore, the random forests was chosen to be the classifier that was used for classifying the identified potential change features.

Random Forest Classifier		
Number of trees	Forest	Non-forest
100	33382	39436
200	33381	39437
300	33381	39437
400	33381	39437
500	33381	39437
600	33381	39437
700	33381	39437
800	33382	39436
900	33382	39436
1000	33381	39437
1500	33382	39436
2000	33382	39436
2500	33381	39437
3000	33381	39437
3500	33381	39437
4000	33382	39436
4500	33382	39436
5000	33382	39436

A

Extreme Random Forest Classifier		
Number of trees	Forest	Non-forest
100	33353	39465
200	33362	39456
300	33351	39467
400	33353	39465
500	33358	39460
600	33355	39463
700	33352	39466
800	33350	39468
900	33355	39463
1000	33355	39463
1500	33353	39465
2000	33352	39466
2500	33350	39468
3000	33356	39462
3500	33352	39466
4000	33352	39466
4500	33353	39465
5000	33353	39465

B

KNN Classifier		
Number of neighbors	Forest	Non-forest
2	33092	39726
4	33012	39806
6	32985	39833
8	32967	39851
10	32957	39861
12	32917	39901
14	32872	39946
16	32859	39959
18	32867	39951
20	32834	39984
22	32825	39993
24	32817	40001
26	32820	39998
28	32805	40013
30	32778	40040

C

Table 6.4: A test of the suitability and stability of the three classifiers over Liwale environment, using different number of trees for random forests and extreme random forests classifiers, and different number of neighbours for K-NN.

Training Data and Parameters for Classifier

Following the identification of potential change features, the non-change features were used to train the classifier to classify the potential change features. To do this, a 5% sample was selected using a histogram sampling approach (implemented within RSGISLib; (Bunting et al., 2014a)). To highlight was that the sample was excluded, and not used later during the validation process. The out-of-bag method was used during the growing of the trees. The classifier was implemented in Scikit learn, a python library that provide state-of-the-art implementations of many well known machine learning algorithms, while maintaining a simple interface that is closely integrated with the Python language (Pedregosa et al., 2011). As stated, the two primary parameters that were required were the number of trees to be built in the forest and the number of possible splitting variables/predictors to be considered for each node in the trees (Prasad et al., 2006; Ismail et al., 2010; Naidoo et al., 2012). The default value for the number of trees value that studies in the southern African environment have used is 500 trees, and it has been reported to produce acceptable results (Ismail et al., 2010; Naidoo et al., 2012). Furthermore to determine the number of possible splitting variables for each node, a standard rule of thumb which is basically the squared root of the total number of predictors is used (Horning, 2010; Naidoo et al., 2012). The number of these randomly selected predictor variables can either be set by the user or the choice can be left to the random forests algorithm (Breiman, 2001; Horning, 2010).

Past studies have concluded that the use of the default 500 number of trees and the use of the standard rule of thumb to determine the number of predictor variables produced acceptable results (Liaw and Wiener, 2002; Horning, 2010; Ismail et al., 2010; Naidoo et al., 2012). Breiman (2001) and Duro et al. (2012) noted that values larger than the default are known to have little influence on the overall classification accuracy. In the Scikit learn random forests module, two options are implemented that can be used for determining the number of splitting variables, and these are the gini index and the entropy index (Salford Systems, 2014), and according to (Pedregosa et al., 2011) neither of these two is better than the other, but are data dependent. This work used the gini index to determine the number of splitting variables for the rule of thumb.

To ascertain if the default number of trees was adequate for the savanna forests of Liwale. First, a 5% training data was sampled across the whole data range. From the remainder of the data (95% non-training data), another 5% was sampled, and the random forests classifier run on it to perform a classification for change across the range of 5 to 5000 trees, at different intervals as

shown in Table 6.5. The results for the different number of trees on the same dataset shows that for these savanna environments of Liwale, the accuracy of the classifier is not affected or dependent on the number of trees. The results show that an increase/decrease in the number of trees does not improve or reduce the accuracy of the classifier. Even though the default 500 trees would have done just as well, a number of 1000 trees (shown in bold in Table 6.5) was used to run the classification as it was found that the training time for the classifier was insignificant. Using a number of 1000 trees and the default gini index, a change classification was performed for the savanna forests of Liwale to produce a change map.

Table 6.5: Parameterisation of random forests classifier for the number of trees.

Number of trees	**Forest**		**Non-forest**	
	Count	*Percentage*	*Count*	*Percentage*
5	33380	99.997	39438	99.997
10	33381	100	39437	100
20	33381	100	39437	100
30	33381	100	39437	100
40	33382	99.997	39436	99.997
50	33381	100	39437	100
60	33382	99.997	39436	99.997
70	33382	99.997	39436	99.997
80	33381	100	39437	100
90	33382	99.997	39436	99.997
100	33382	99.997	39436	99.997
200	33381	100	39437	100
300	33381	100	39437	100
400	33381	100	39437	100
500	33381	100	39437	100
600	33381	100	39437	100
700	33381	100	39437	100
800	33382	99.997	39436	99.997
900	33382	99.997	39436	99.997
1000	**33381**	**100**	**39437**	**100**
1500	33382	99.997	39436	99.997
2000	33382	99.997	39436	99.997
2500	33381	100	39437	100
3000	33381	100	39437	100
3500	33381	100	39437	100
4000	33382	99.997	39436	99.997
4500	33382	99.997	39436	99.997
5000	33382	99.997	39436	99.997

6.4 Results

This section presents the change results for Liwale for the period 2012-2014 that were produced using the random forests classifier, based on potential change features that were identified using the 2012 forest baseline and the 2014 wet season RapidEye image. This change map product was compared to the 2014 LiDAR data over the SLU site in Liwale. The product was further validated for the whole of Liwale using a set of sampled points across the site and reference images.

6.4.1 Change Map from RapidEye Images

The classification baseline-to-image change detection approach produced a forest change product for Liwale between May 2012 and February 2014. Figure 6.9 is the 2014 change map of the whole site of Liwale displayed against the forest baseline of 2012. As shown in the change map where change areas are shown in red, it was found that the part of the site experiencing forest disturbance most was the south-eastern location where there is currently a lot of clearing. Most of this clearing is primarily for agricultural purposes whereby the practice of shifting agriculture at subsistence level is rife.

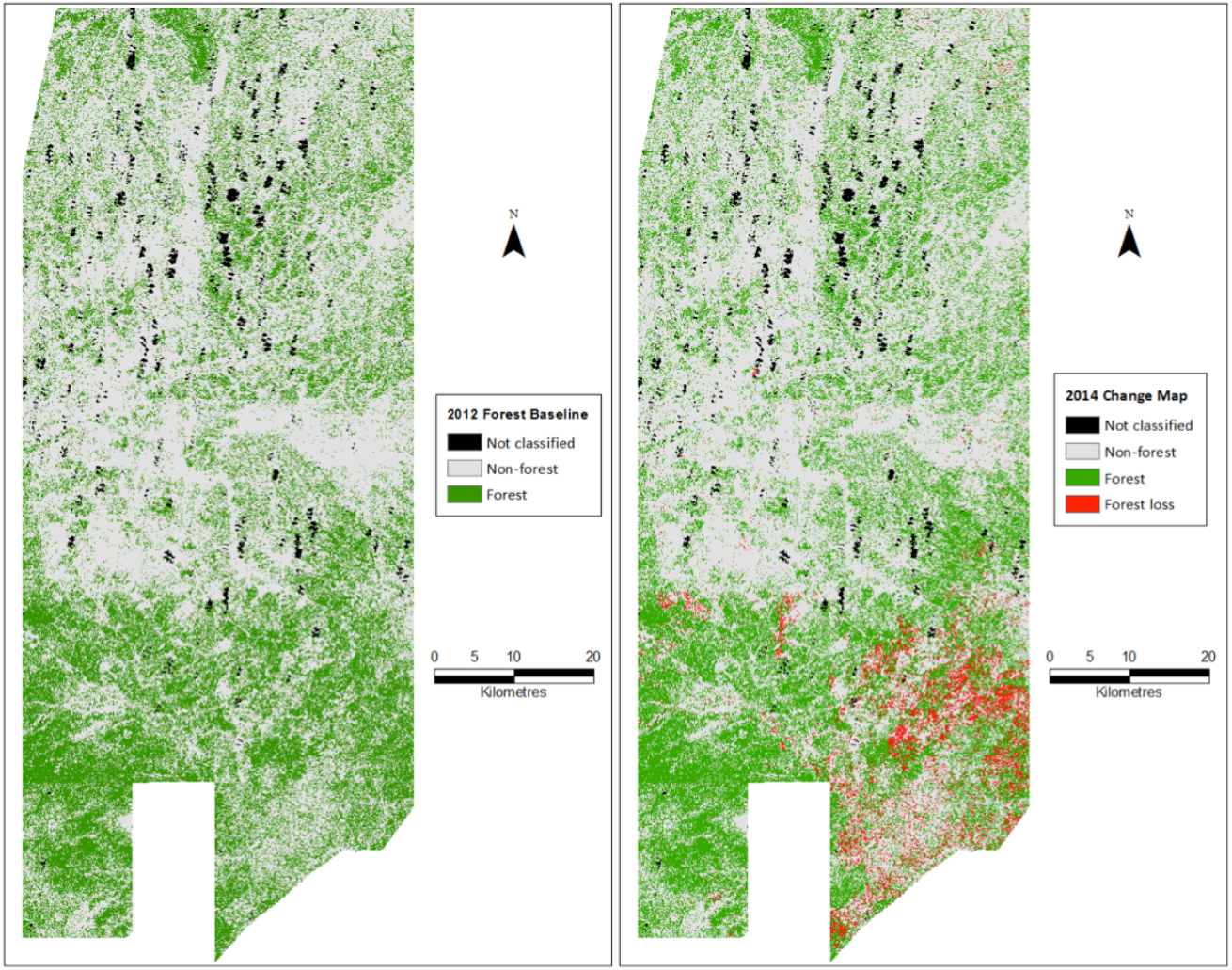


Figure 6.9: The change map product of Liwale produced from the 2012 forest baseline and a 2014 wet season imagery.

Table 6.6 is a summary of land cover statistics for the change map of the site of Liwale. Liwale lost approximately 15550 hectares (155.5 km^2) of forest, a 9% loss within the two years. As described above, this was mainly concentrated in the south-eastern part of the site where there is a lot of settlements and dominated by shifting agriculture practices. The non-forest class covered 3391.39 km^2, while the forest class saw a decrease from 1779.82 km^2 to 1624.32 km^2.

Table 6.6: The change statistics between the years 2012 and 2014 in Liwale.

Class	**Area** (km^2)	
	2012 Forest baseline	*2014 Change Map*
Not classified	110.53	110.53
Non-forest	3391.39	3391.39
Forest	1779.82	1624.32
Forest loss		**155.50**
Total (km^2)	5281.74	5281.74

Figure 6.10 shows an example of the false colour images of the two years 2012 and 2014 and the corresponding forest baseline and change product respectively of the areas around the SLU site. The area has been worst affected by disturbance between 2012 and 2014, as shown in the map. In 2012, over 50% of the area was forest. Two years later, a better part of the SLU, from its centre to the eastern part was heavily disturbed for crop production.

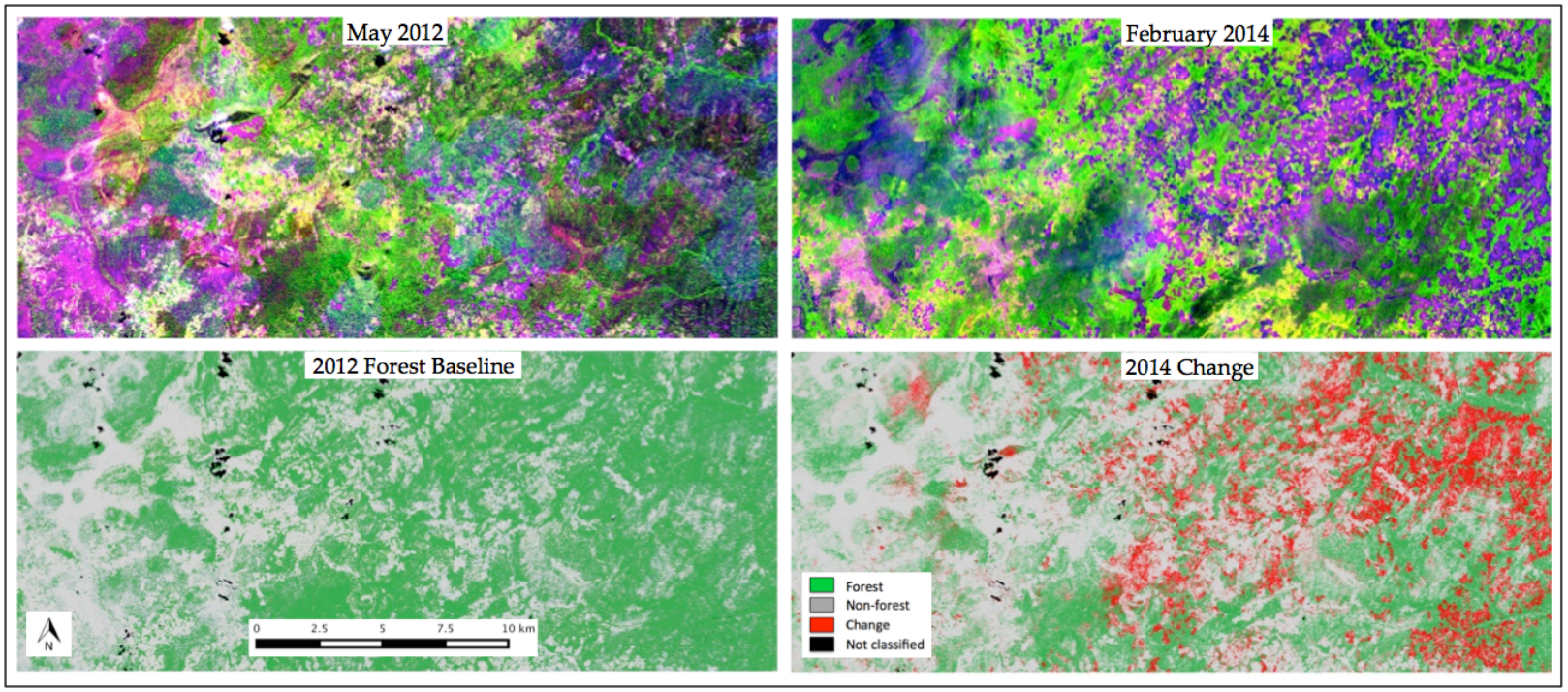


Figure 6.10: An area within the Liwale site known as the SLU where there is a lot of forest disturbance resulting mainly form forest clearing as a result of shifting cultivation. The RapidEye false colour image was displayed as B:RedEdge, G:NIR, R:Blue.

Most forest mapping techniques developed were mainly focused on large scale forest disturbance, and this is relatively easier to map. The African environments and specifically the Tanzanian environment, as noted under literature review in Chapter 2, experiences both large-scale (deforestation) and small-scale (degradation) disturbance. The method that was developed for detecting change therefore was one hoped to be able to pick up both scales of change. Figure 6.11 is a subset of an area where there was substantial large-scale disturbance for agricultural practice. For large-scale clearing in the environment, the February 2014 false colour image in Figure 6.11 shows the distinct difference between the forest areas and the non-forest areas with the latter being inclusive of the change areas. The distinct spectral difference in the change areas from the forest class made the method perform well in mapping change at this scale.

The technique was found to also effectively map small-scale change, as illustrated in Figure 6.12. Shown with the red polygon is an area that was forest in 2012, as seen in the 2012 false colour image. However, the 2014 false colour image shows a change in the spectral signature of the pixels, depicting disturbance. The technique picked the pixels within the polygon as change. However, the main limitation with the mapping of very small-scale change whereby it covered only a few pixels was the noted poor registration of the RapidEye data. A shift between the images of the respective years being used to establish the forest baseline, and then the detection of change, even if as low as a pixel, is likely to affect the detection of change, and thus reduce its ability and accuracy.

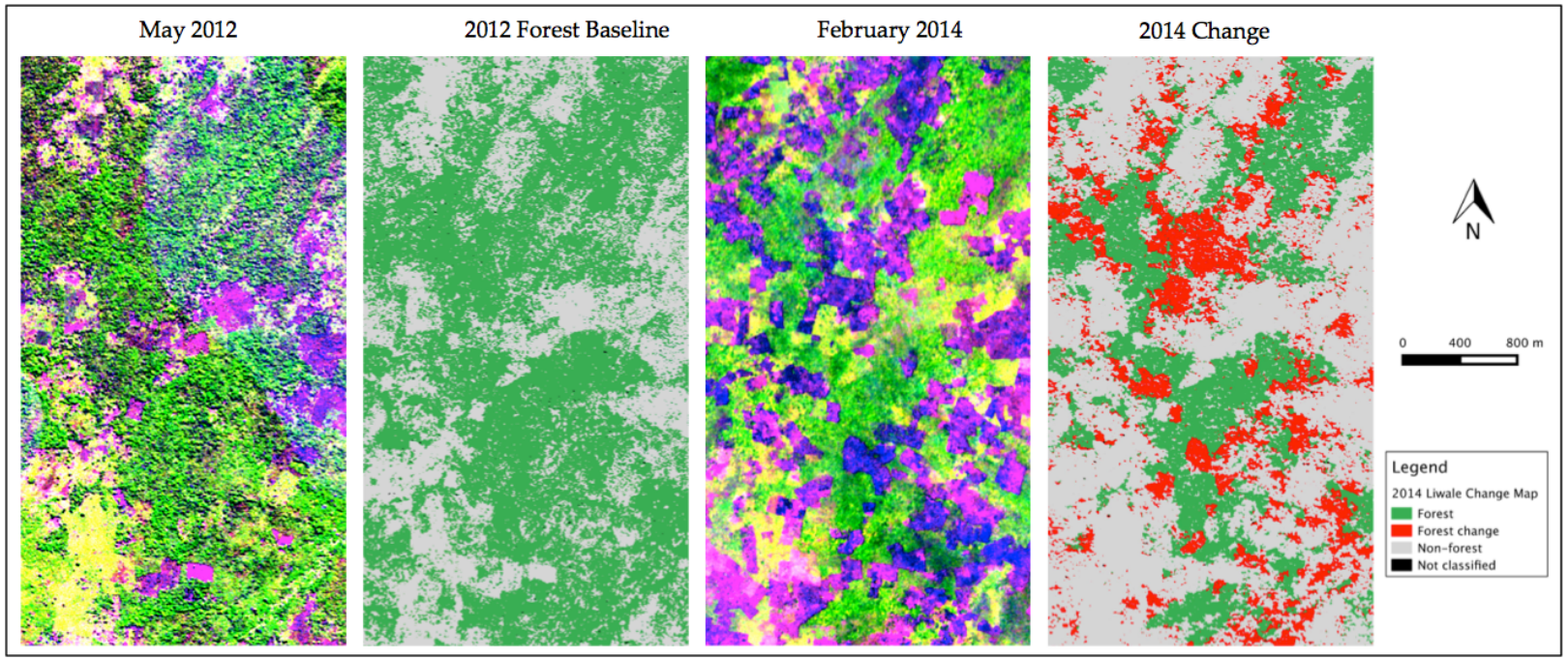


Figure 6.11: Change detection over an area experiencing a lot of shifting cultivation and clearing at large-scale level. The RapidEye false colour image was displayed as B:RedEdge, G:NIR, R:Blue.

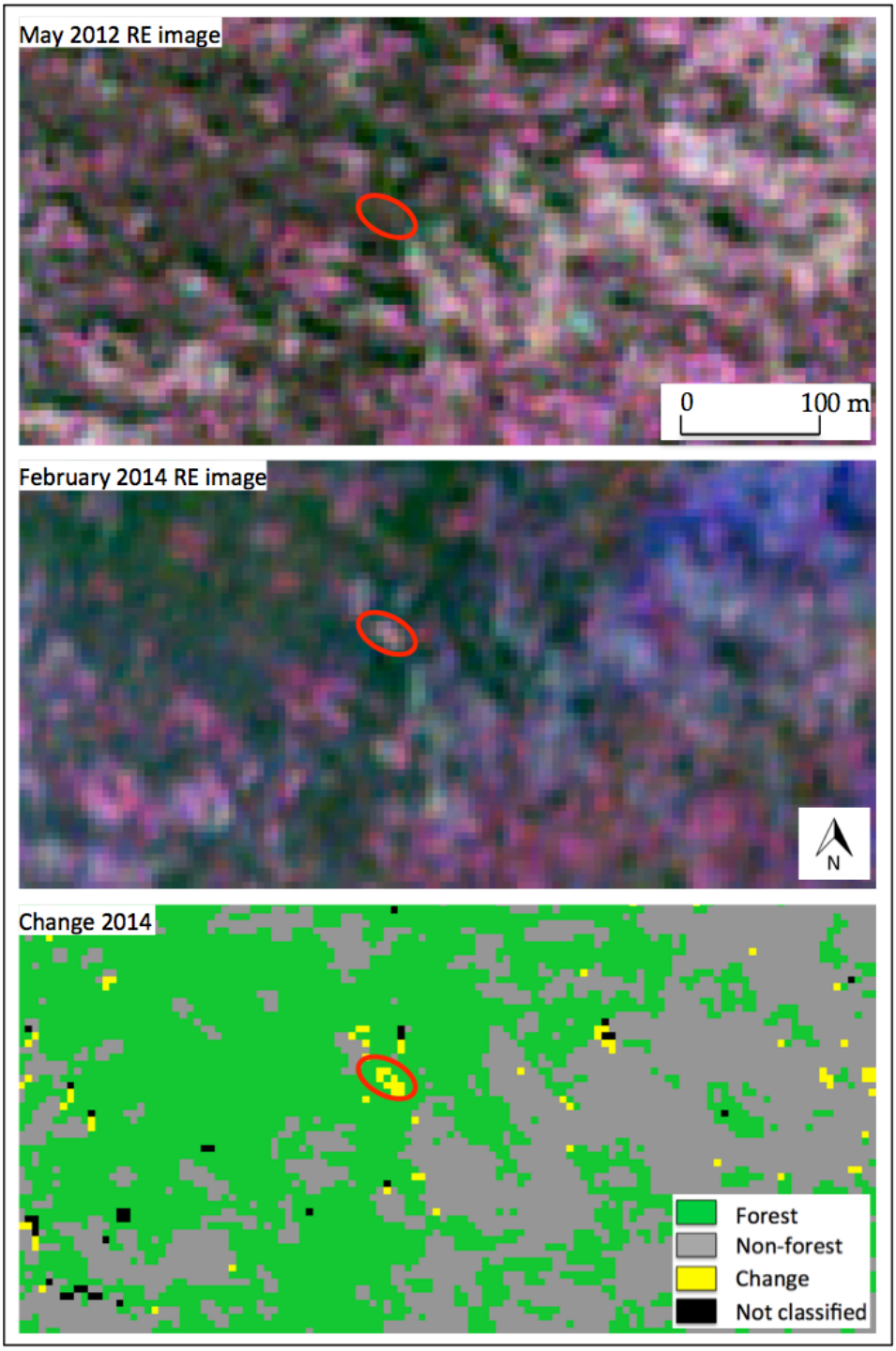


Figure 6.12: An illustration where the change detection technique has correctly captured a small change feature. The RapidEye false colour image was displayed as B:NIR, G:RedEdge, R:Blue.

6.4.2 2014 Updated Forest Baseline for Liwale

From the 2014 change map product, the non-forest and forest loss classes were merged into one non-forest class, and in the process produced an updated forest/non-forest baseline map of Liwale for the year 2014. Figure 6.13 is the updated forest baseline for the whole site together with the forest baseline of the year 2012. As discussed, the baseline mainly changed on the southern part of the site where there is relatively closed forests and a lot of human settlements. For a small scale change map, the extent and degree of change does not show clearly, probably due to that the change areas are small and fragmented. However, a zoom into the areas experiencing disturbance at a higher rate and scale in southern Liwale as depicted in Figure 6.14, it can be seen in that an extensive amount of forest had been lost between 2012 and 2014.

Figure 6.13: The updated 2014 forest baseline of the site of Liwale produced from the 2014 change map of the area.

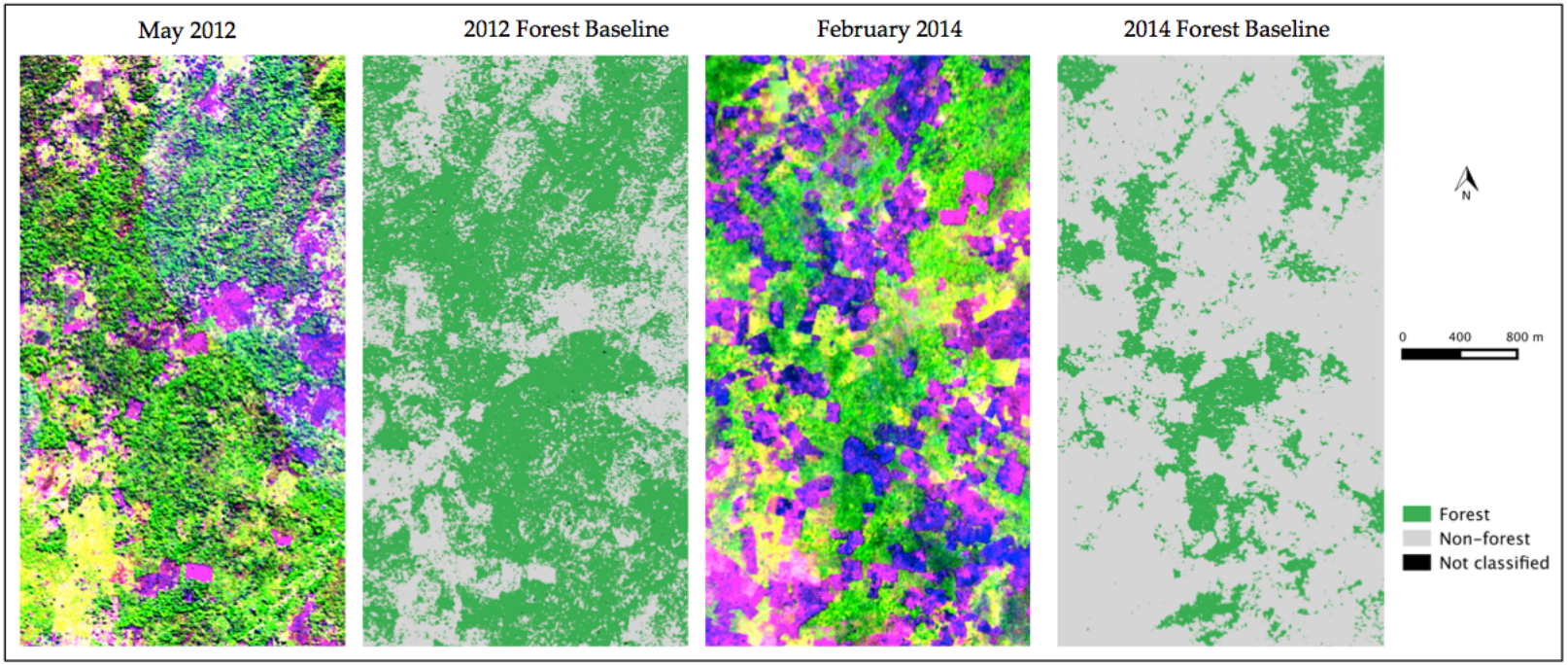


Figure 6.14: An updated forest baseline for the year 2014 from the 2014 change map. The RapidEye false colour image was displayed as B:RedEdge, G:NIR, R:Blue.

Table 6.7 shows that considering the same area between the two years, the forest area was reduced from 1779.82 km^2 to 1624.32 km^2, which showed a loss of 8.74%. An increase of 4.59% in non-forest areas saw the class coverage increase from 3391.39 km^2 to 3546.89 km^2 between 2012 and 2014.

Table 6.7: The updated 2014 forest baseline of Liwale versus the 2012 forest baseline.

Class	**Area**		
	2012 Forest Baseline (km^2)	*2014 Forest Baseline (km^2)*	*Loss/Gain (%)*
Not classified	110.53	110.53	
Non-forest	3391.39	3546.89	+ 4.59
Forest	1779.82	1624.32	- 8.74
Total	5281.74	5281.74	

6.4.3 Comparison of Updated Forest Baseline to the LiDAR Data

Having shown that the proposed REDD forest definition of Liwale produced results with a better accuracy than the currently adopted forest definition, the comparison between the produced change product and the LiDAR data was performed only on the CHM meeting the attributes of the proposed definition. The comparison was undertaken using 30 m, 100 m and 250 m grids, corresponding with the analysis undertaken in Chapter 5 for the classifications.

The overall accuracies for the comparison over the 30 m, 100 m and 250 m grids were 79.46%, 80.9% and 81.06% respectively. For the user accuracy, both forest and non-forest classes had high accuracies of over 78% for all three grids. The producers accuracy for the forest class was high for all three grids, yet relatively lower for the non-forest class. It was less than 70% for all three grids, and notably was 67% for the 30 m and 100 m grids, but dropped to 63% for the 250 m grid.

Table 6.8: Comparison of the change product to the CHM derived from the LiDAR data over a 30 m grid.

	Forest	*Non-forest*	*User*	**User (%)**
Forest	37351	8426	45777	81.59
Non-forest	18290	66027	84317	78.31
Producer	55641	74453	130094	
Producer (%)	67.13	88.68		**79.46**

Table 6.9: Comparison of the change product to the CHM derived from the LiDAR data over a 100 m grid.

	Forest	*Non-forest*	*User*	**User (%)**
Forest	3345	564	3909	85.57
Non-forest	1637	5976	7613	78.50
Producer	4982	6540	11522	
Producer (%)	67.14	91.38		**80.90**

Table 6.10: Comparison of the change product to the CHM derived from the LiDAR data over a 250 m grid.

	Forest	*Non-forest*	*User*	**User (%)**
Forest	486	68	554	87.73
Non-forest	276	986	1262	78.13
Producer	762	1054	1816	
Producer (%)	63.78	93.55		**81.06**

6.4.4 Validation of Change Map Using Random Points

Lawrence et al. (2006) and Naidoo et al. (2012) are among studies that suggest that when random forests is used, there may be no need for cross validation or a separate test dataset to determine any misclassification error because the OOB error provides an unbiased estimate of error. However, other studies concluded that using the OOB error to determine the misclassification error could result in a biased estimation of the error as the samples used to calculate the error are not independent of the model being evaluated (Díaz-Uriarte and Alvarez de Andrés, 2006; Granitto et al., 2006). Ismail et al. (2010) avoided bias in the accuracy assessment by using an independent test dataset of 390 points, tabulating the results using a confusion matrix. This work used a similar approach whereby an independent validation process was performed using a set of sample points taken across the study site, and the change product was validated using reference images (RapidEye and Google Earth).

The Validation Process

From the derived forest baseline, a forest binary mask was created. This binary image was then used to clip out all objects in the 2014 change map that were forest in 2012. On the resultant clipped image, an object was either forest or change (forest areas that had experienced disturbance).

Using this new (clipped) image, stratified random sampling was performed to create a shapefile of a set of points for both the forest class and the change class, and these sampled points were populated with the corresponding 2014 class in the attribute table. On the same image, transects were manually digitized on the boundaries of the forest and change areas in ArcMap software. Then along the transects, points that were 20 m apart were generated. This resulted in another shapefile with a set of validation points for both the forest and change classes, and each point was populated with the class type for 2014.

The stratified randomly sampled points and the transect points shapefiles were then loaded into QGIS software, and a 5 m buffer was created around each point. As discussed in Chapter 5, this ensured that a window of four pixels was considered at each instance, than an individual point. Then using the accuracy assessment tool in QGIS, outlined in Chapter 5 (Figure 5.22), these sampled points were validated. The validation was based on a 2014 false colour RapidEye composite image which was used as the primary reference image, and a Google Earth image which was used as a secondary reference image. The reasons why Google Earth was not used as the primary reference image or in singularity as a reference image are discussed in Chapter 5, under the validation section. Primarily, the reasons boiled down to poor coverage over Liwale by Google Earth, and Figure 5.20 and Figure 5.21 illustrate some of the challenges. These reasons still necessitated the use of the RapidEye composite as the primary validation image, and the Google Earth image as a secondary validation image at this stage.

As highlighted in Chapter 5, Olofsson et al. (2013) recommends that further statistical analysis be undertaken to validate the accuracy assessment for land use maps, and preferably not use kappa coefficient. The paper noted that even though the overall accuracy and kappa coefficient measures are commonly published, these were insufficient to fully address the area estimation and uncertainty information needs. Using the statistical approach explained in Olofsson et al. (2013) and Olofsson et al. (2014), accuracy measures were further presented with 95 % confidence interval. The kappa coefficient was only used as a supporting statistic.

In addition, Olofsson et al. (2013) and Olofsson et al. (2014) presented an approach for estimating the error-adjusted land change area, with a 95% confidence interval, which was noted to yield valuable information on change that may be very different from the results obtained solely from the map. This is a very useful method that avails an opportunity to present the area of change values for the classes of interest with a level of uncertainty.

Because of classification error, the mapped area proportions are usually biased when the objective is to estimate the true proportion of area of a category as determined from the reference classification. Instead of obtaining the area directly from the map classification, the method bases an area estimator on the reference classification of each sample unit. The area proportions for each reference-defined category are then estimated, and in turn, an unbiased estimator of the total area (based on the reference classification) of each category is estimated (Olofsson et al., 2014).

This stratified estimator, which was used in this study as a post-stratified area estimator (Olofsson et al., 2013), is the error-adjusted estimator of area as it includes the area of map omission error of a category and leaves out the area of map commission error. The estimated standard error of the estimated area proportion is then computed as detailed in Cochran (1977). Thereafter, the standard error of the error-adjusted estimated area is calculated, and the approximate 95% confidence interval derived. The estimated total area of forest loss was therefore, further adjusted for error using the formulas presented in Olofsson et al. (2013).

Validation Results

From the stratified random sampling process, 3000 points were generated. Then along transects sampled across class boundaries, a total sample of 4554 points was generated. The sample of 3000 points sampled through stratified random sampling was classified for change with an overall accuracy of 96.8% (kappa coefficient of 0.942). For the transects points sample, it was found that the overall change detection accuracy was 96.75%, with a kappa coefficient of 0.932. Against the thought that along class boundaries, the change detection would be prone to error, the results showed that the technique was capable of correctly demarcating the boundaries and correctly classifying for change as the overall accuracy was the same for both set of points (Table 6.11). For both set of samples, the user and producer accuracies were found to be greater than 95% for both the forest and non-forest classes. accuracy

Table 6.11: Overall accuracies of the set of stratified random points, transect boundaries points and the combined points.

Set of points	**Overall accuracy**	**Kappa coefficient**
Random points	96.8	0.942
Transect points	96.75	0.932
Combined points	96.77	0.943

An overall accuracy for the change detection in Liwale was then computed by combining the set of sample points, and Table 6.12 is the error matrix for the combined points. The overall accuracy of detecting change in Liwale was 96.77%, with a kappa coefficient of 0.943, which shows that the method was very effective for change detection in the savanna environment of Liwale. As was the case for the two sets of points, the user and producer accuracies for both the forest and non-forest classes for the combined set of points was found to be greater than 96%, an indication of a high level of accuracy.

Table 6.12: Error matrix for the change map based on all 7554 points generated by stratified random sampling and along transects in forest and change boundaries, across the image. Map categories are the rows while the reference categories are the columns.

	Forest	*Forest loss*	*User*	**User (%)**
Forest	3737	100	3837	97.39
Forest loss	144	3573	3717	96.13
Producer	3881	3673	7310	
Producer (%)	96.29	97.28	96.77	**96.77**

For the combined random points, accuracy measures were then further presented with 95% confidence interval, to fully address the area estimation and uncertainty information needs (Olofsson et al., 2013). Table 6.13 presents the user's and producer's accuracies for the change map, and the overall accuracy of the change map was 0.97 ± 0.005. The change map had user's accuracies of 97% for the forest class and 96% for the forest loss class, with uncertainties of ±0.5% and ±0.6% respectively. Notably, these accuracies were similar to those derived from the use of an error matrix. However, for the producer's accuracies, it improved to 100% for the forest class, with an uncertainty of ±0.1%. For the forest loss class, the producer accuracy decreased to 78%, with an uncertainty of ±4.3%.

Table 6.13: Error matrix for the 2012 to 2014 period change map of Liwale. Map categories are the rows while the reference categories are the columns. Accuracy measures are presented with a 95% confidence interval and the overall accuracy was 0.97 ± 0.005.

Class	***Forest***	***Forest loss***	***Total***	***Am [sq. km]***	***Wh***	**User's**	**Producer's**
Forest	3737	100	3837	3547	0.686	0.97±0.005	1.00±0.001
Forest loss	144	3537	3681	1624	0.314	0.96±0.006	0.78±0.043
Total	3881	3637	7518	5171	1		

Table 6.7 shows that the estimated forest loss between 2012 and 2014 in Liwale was 155.5 km^2. Using the error-adjusted estimator (Olofsson et al., 2013), the stratified estimator of the disturbed

forest area was found to be 191.75 km^2. The standard error of the adjusted area estimate was found to be 31.90 km^2. This gave a final land change area estimate with a margin of error (at approximate 95% confidence interval) of 191.75 $\pm$ 63.80 km^2. Taking this uncertainty into account, the true area of forest loss could be as low as 127.95 km^2 or as high as 255.55 km^2 at the 95% level of confidence.

6.5 Discussion

6.5.1 The Change Detection Approach

The research developed a technique that uses an existing forest baseline and a newly acquired satellite image to detect change in the savanna forests of Tanzania. This is a crucial component of a simple monitoring system that would be pivotal in the sustainable management and utilisation of the forest resources, and especially keeping track of changes that take place in them, both their scale and direction.

Overall, the change detection method proved to hold a great potential for detecting change in the Tanzanian savanna forests. The method yielded an overall accuracy of 96.77%, with a kappa coefficient of 0.943, which shows that it worked well for the environment. It was tested and found to be robust, as well as possessing the ability to be transferred across to the wider savanna environment of Tanzania and southern Africa. The Liwale and the Tanzanian environment at large, experience both deforestation and forest degradation, with the latter as discussed in Chapter 2, being the harder to map. The method was found to do well in capturing/mapping small change features as well, meaning that both degradation and deforestation can be accounted for.

To support the accuracy statements, further statistical analysis was undertaken, which according to Olofsson et al. (2013), fully addresses the area estimation and uncertainty information needs. First, the accuracy measures were presented with a 95% confidence interval, which factored in the areal coverage of each class. The overall accuracy derived from this statistical measure for a 95% confidence interval was 97%, with $\pm$ 0.5% uncertainty. The measure highlighted that there was actually much higher omission in the classification for forest loss in Liwale, with the producer's accuracy for the class decreasing to 78%, compared to 97% derived from the image-based error matrix. However, the user's accuracies for both forest and forest classes were found to be similar

to those derived from the image-based error matrix. This was also consistent for the producer's accuracy of the forest class.

Secondly, the stratified estimator was used as a post-stratified area estimator (Olofsson et al., 2014) to adjust for error in area estimates in the change map, and thus derive a better estimate of forest loss in Liwale between 2012 and 2014, with a level of uncertainty. The mapped area for forest loss was 155.5 km^2 and the stratified estimator of the disturbed forest area was found to be 191.75 km^2, which shows that there was a difference of 36.25 km^2 on the two estimates. This difference can be explained as coming from the fact that the forest loss class was classified with a producer's accuracy of 78%. This means that 22% of the proportion of area under forest loss class was omitted from the map. This led to the discrepancy since the error-adjusted estimate of forest loss area adds even those omitted areas to the mapped area of the forest loss during the adjustment process. With a standard error of the adjusted area estimate being 31.90 km^2, the final land change area estimate with a margin of error (at approximate 95% confidence interval) was 191.75 $\pm$ 63.80 km^2. This uncertainty means that the true area of forest loss could be as low as 127.95 km^2 or as high as 255.55 km^2 at the 95% level of confidence.

Possible alternatives for interrogating the change map included comparing it with Hansen et al. (2013), which is a global forest change product. However, there main limitation with the approach was that the product was for the period 2000-2012, yet the produced map was for 2012-2014. The different time series was a limitation, especially having shown using LiDAR data in Chapter 5, forest change (loss) taking place even within a space of three months in the environment. Moreover, being a global product with a relatively higher resolution than the product that had been produced from RapidEye, the products were found to differ greatly in terms of forest/non-forest coverage. For this reasons, the two products could not be compared.

A monitoring system in place would mean that the next available image that is appropriate for change detection would be compared to this updated baseline, and thus serve to show the pattern and direction of forest change between the year 2014 and whatever year it is acquired. From the resultant new change map, an updated baseline would be produced. The monitoring system would accumulate a time series of forest baselines. The more the time series data accumulation, the easier it will be to correct for errors in earlier baseline classifications.

6.5.2 RapidEye for Change Detection in Liwale

RapidEye imagery was found to be effective in capturing both small- and large-scale change in the savanna woodlands. The imagery gives adequate spatial and spectral detail to enhance the detection of disturbed (degraded and deforested) areas. However, it was also noted that its limitation lies in that it is insufficient to resolve individual features such as trees, yet it also captures too much of the detail within forested landscapes, which compromises the delineation of blocks of forest and change areas.

Forest disturbance was not randomly spread out across the site of Liwale, rather it was more dominant and pronounced in settled areas or in the vicinity of human settlements. The northern part of the site is the Selous Game Reserve, which is a protected site, and thus with minimal human interaction with the forests there. The small scale disturbance picked in these protected areas is mainly as a result of forest fires, which as shown in the fieldwork section in Chapter 4, can be very intense to the point of burning down even big trees. Moreover, during the fieldwork, it was observed that the large herbivory browsers such as the elephants sometimes destroyed trees to the point of either ripping them or flooring them. Some parts of these savanna were also notably of poor soils such that trees were toppled by winds as well. These are amongst the minimal disturbances picked in the protected areas.

Other protected areas in Liwale include the Angai forest reserve, which was gazetted by different communities. These reserves have also remained well protected, and experience minimal clearing compared to nearby forest areas to which communities can expand or shift into. In some areas, there is a clear demarcation between the protected and unprotected areas, with the forests in the unprotected areas exhibiting heavy disturbance and a sharp contrast to the protected ones.

Worth noting though is that the same limitations the imagery had in the establishment of the forest baseline would still affect the change detection process. The lack of a strong time series data, and ideally a number of good images for the same season limits the ability of producing a change map with a very high overall accuracy. As discussed in Chapter 4, seasonality and persistent cloud cover greatly reduced the availability of data appropriate for forest mapping and change detection. In years where there were more image acquisitions, seasonality made most of the data inappropriate.

In the discussion section in Chapter 5, it was highlighted how the high level of heterogeneity in

these environments lead to an overlap in the feature space, which in turn introduces complexity in the separation of the features. It was highlighted that even though it is a high resolution image (5 m), it is a spectrally poor image that is not capable of showing the boundaries of the classes. Poor segments with overlaps to other classes result in a baseline with those overlaps incorrectly classified, and that error propagating into the change detection. Even though the assessment of how well boundaries of change features and the forest showed that these had been well classified, boundaries were hard to determine.

Another limitation lied in the poor registration of RapidEye in the region, which Black Bridge is working towards rectifying (RapidEye, 2012). Figure 5.21 is an example of how poorly registered the dataset may be, as seen on the Google Earth image where it was noted to be off compared to other datasets.

6.5.3 Sources of Error in the Change Detection and Validation

The noted slight reduction in producers accuracy for the non-forest class in the comparison of the updated 2014 forest baseline derived from RapidEye data to the 2014 LiDAR data implies there was an overestimation of change class, and thus the non-forest class. The major sources of error discussed in Chapter 5 still held true in the change detection level. In addition, having used the forest baseline which had an overall accuracy of 0.94 ± 0.012 meant that the error represented by the $\pm$ 1.2% uncertainty on the baseline was propagated into the change product during the change detection process. The noted shift in the RapidEye datasets, yet this is a highly fragmented environment is fragmented, with a lot of gaps within the forest class, means that there are some areas or part of change areas that are incorrectly captured as change due to the shift.

Furthermore, as discussed above, the same errors experienced during the establishment of a forest baseline were noted to affect the overall accuracy of the change map. The complexity and high heterogeneity of the environment leads to a lot of overlap in the feature space. The unsystematic shift in the RapidEye imagery, which was reported to be up to 5 m in other other studies would still be a challenge for the change detection, especially due to the fragmentation of the environment. The comparison with LiDAR data which had detail so as to model even features made up of just an individual or a couple of pixels, a shift may have a large effect. As in Chapter 5, the limitation introduced by the shift was compensated for by undertaking the comparison between the change product and the LiDAR data at grid level, instead of at feature level. However, error

compensation works best for not heavily and heterogeneously fragmented environments. The shift effect was therefore compensated in both the RapidEye base data and the RapidEye and Google Earth reference images by using buffered sample points, than in singularity. That way, it ensured that validation was performed over a window of four pixels than a single pixel. The difference between the acquisition times of the images used and the validation (base) images was likely to introduce error, as shown in Chapter 5 where the environment had experienced relative change in a space of three months within the wet season (Figure 5.23)

6.6 Conclusions

A change detection method was developed for the savanna forests of Liwale. It uses an existing forest baseline and a new RapidEye image to identify areas of potential change, which are then further classified into either change or non-change using random forests classifier, and thus produce a change map. An updated forest baseline can then be produced from the change map. The technique was found to identify and successfully map change features with an overall accuracy of 96.77% and a kappa coefficient of 0.943 in the savanna forests of Liwale. When presented with 95% confidence interval, the overall accuracy was 97%, with an uncertainty of $\pm$ 0.5%. Using the post-stratified area estimator, the standard error of the adjusted area estimate was 31.90 km^2, and the final land change area estimate with a margin of error (at approximate 95% confidence interval) was 191.75 $\pm$ 63.80 km^2. The uncertainty means that the true area of forest loss in Liwale could be as low as 127.95 km^2 or as high as 255.55 km^2 at the 95% level of confidence.

This method, building up from the forest baseline establishment method (Chapter 5), is robust and has the capability to be transferred across to the wider savanna environment of Tanzania and Africa. Forest/non-forest change detection is a critical component of a sound and successful forest monitoring for the savanna forests, and this approach holds a great potential, especially because it was found to be capable of mapping small change features resulting from forest degradation, a practice common in these environments. However, the limitations resulting from data unavailability due to seasonality and persistent cloud cover over Tanzania were found to hold true for forest monitoring in Liwale and Tanzania.

Chapter 7

Scaling Methods to Landsat Imagery

7.1 Introduction

High resolution remote sensing data are typically costly and country-wide acquisitions are often difficult to justify. There are a few countries (e.g., Guyana) where support has been provided, which allows them to fund companies (e.g., Black Bridge) to task imagery. However, such funding and programs are not open to all countries as some may not be able to fund companies to task imagery. The cost of imagery also increases where cloud-cover is prevalent, and in Tanzania the challenge of cloud cover has been discussed at length in earlier chapters. Moreover, the cost of processing these data to a level that is usable (e.g., for REDD+ activities) can be high. The expense associated with obtaining and processing lower resolution imagery (e.g., Landsat and Sentinel-2) is much lower by comparison, although cloud-cover still persists, and the resolution may be insufficient to resolve the processes of forest disturbance, especially degradation. Therefore, a potential solution is to use a combination and/or integration of both VHR, HR and MR imagery. The following sections highlight this potential by showing how a low resolution image can be used as well for both forest mapping and monitoring. This was done by scaling out the techniques that had been developed using HR RapidEye imagery to the low resolution Landsat imagery.

The RapidEye sensor acquire data with a spatial resolution of 5 m. A difficulty with this resolution

is that it is insufficient to resolve individual trees yet it also captures too much of the detail within forested landscapes, which compromises the delineation of blocks of forest. By comparison, the Landsat sensor data have been used for LULC classifications and change detection from local to global scales and serves as the key source of data for many countries (Manandhar et al., 2009; Markham and Helder, 2012; Hansen et al., 2013). Key benefits of this sensor is that long-time series of data are freely available at a global level, with a high frequency of observations. The scale and spectral range are also well suited to forest cover mapping (Cunningham, 2014).

7.2 Background

7.2.1 Forest Baseline Classification

At a local scale, field-based ecological surveys usually suffice for mapping change. However, these become expensive and time consuming for regional and greater scales. Healey et al. (2005) observed that feasible means of mapping and monitoring forest change on a regular and continuous basis at large scales is with the aid of remote sensing data. The Landsat sensors have provided useful for short to long-term monitoring of forests (Hansen and Loveland, 2012; Markham and Helder, 2012; Hansen et al., 2013). By providing information on LULC over time and space, these data can contribute significantly to environmental monitoring. This will ensure that necessary actions are taken to limit the impacts of deforestation, including through policy development.

To support the monitoring of global deforestation, Hansen and Loveland (2012) and Hansen et al. (2013) classified time-series of Landsat sensor data within the Google Earth Engine. This was the first attempt at undertaking such an analysis which was supplemented with a map of forest and non-forest for four years (2007-2010) based on time-series ALOS PALSAR data (Shimada and Ohtaki, 2010; Hansen et al., 2013). Yuan et al. (2005) also demonstrated that Landsat sensor data provided a reliable and economical means of LULC monitoring. Overall accuracies for a seven-class image classification averaged 94% over a four year period. Accuracies for land cover change were ranged between 80% and 90%. However, the accuracy of any change maps is dependent upon the accuracy of the individual classifications and errors are also propagated.

7.2.2 Forest Change Detection

Change detection is the process of identifying differences in the state of an object or phenomenon through observing it at different times (Singh, 1989). The detection of forest disturbance is important both in research and policy related to global carbon cycles (Hansen and Loveland, 2012). It is a valuable step that is useful for identifying spatio-temporal trends, and thus inform forest management, both locally and at a global scale.

Remote sensing data continues to be a major data source for change detection studies. Chen et al. (2012) highlights that remotely sensed data is a significant source of data for change detection due to the wide range of spatial and spectral resolutions, high temporal frequency and digital format that makes it suitable for computational processing. Methods and techniques that have been developed and documented on the use of remote sensing for change detection have been reviewed above in Chapter 6, and are documented in Coppin et al. (2004), Lu et al. (2004), Alqurashi and Kumar (2013), and Hussain et al. (2013). The techniques use multi-temporal RS datasets to undertake a qualitative analysis of the temporal effect of the phenomena, and then quantify the changes (Hussain et al., 2013). The underlying principle of these methods is that they analyse the process and trend of changes by monitoring ground objects based on time series (continuous) remote sensing observation data (Alphan, 2011).

This study sought to scale the change detection that had been developed for the savanna forest of Liwale, using RapidEye imagery, to the lower resolution Landsat data, to quantify the ability of Landsat data to detect changes in these savanna environments. The advantage with being able to use Landsat data for change detection is that, as mentioned, Landsat has a rich spatial and temporal archive globally, and therefore would be cost-effective for monitoring the Tanzanian savannas.

7.3 Methods

7.3.1 Forest Baseline Classification

To generate a forest baseline from the Landsat sensor data, the same approach applied to the RapidEye data was used. To train the baseline, the map generated from the RapidEye data was used. However, as discussed in Chapter 5, Landsat has a relatively very poor time series coverage

over Liwale. In 2012, the year that coincide with the forest baseline derived from RapidEye, there was no suitable image because of cloud cover, but also differences in the time (and hence season) of acquisition. Hence, data from the 2014 wet season were used as these coincided with the acquisition period of the RapidEye data, and did not have seasonality limitations.

The Landsat image was preprocessed, and then calibrated to allow the calculation of indices, with these normalised, and then stacked with the spectral bands. The resultant image was then segmented and the RAT was populated with image statistics. These statistics were then used to develop rulesets which were used to perform a classification according to the LCCS taxonomy, to produce a forest baseline for 2014. The details of each step of the processing techniques are detailed below, and Figure 7.1 is a schematic overview of the process.

Indices and Image Segmentation

Many of the indices used for the RapidEye classification relied on the red edge band which is not available on the Landsat sensor. Therefore, a number of indices that had been used could not be used, but indices relevant to the spectral bands of Landsat were computed. These were, however, found not to be effective or any better than the spectral bands themselves in the separation of different classes for Landsat. The false colour image of a combination of the NIR, SWIR1 and blue bands was found to show a good separation amongst most classes, and thus those bands together with NDVI were targeted.

An NDVI image was therefore produced for Liwale, and offsetted to be centred on 100 in order to move all values away from the 0 value which is taken as a 'no data' value in RSGISLib. The normalised NDVI product was then stacked into the original image. The resultant stacked was then segmented. As explained in Chapter 5, this algorithm employs two key parameters, namely, the number of clusters (k) in the K-means which are used to seed the K-means algorithm, and the minimum object size in pixels to which objects are eliminated. As with RapidEye, the k value used was 20 and to cater for the high fragmentation in the environment, the minimum object size was kept at 2 pixels. The resultant RAT image was then populated with the statistics of each of the spectral bands and NDVI. The statistics populated into the RAT attribute table for each object were the minimum, maximum, standard deviation and the mean values of that object for each of these bands and the NDVI.

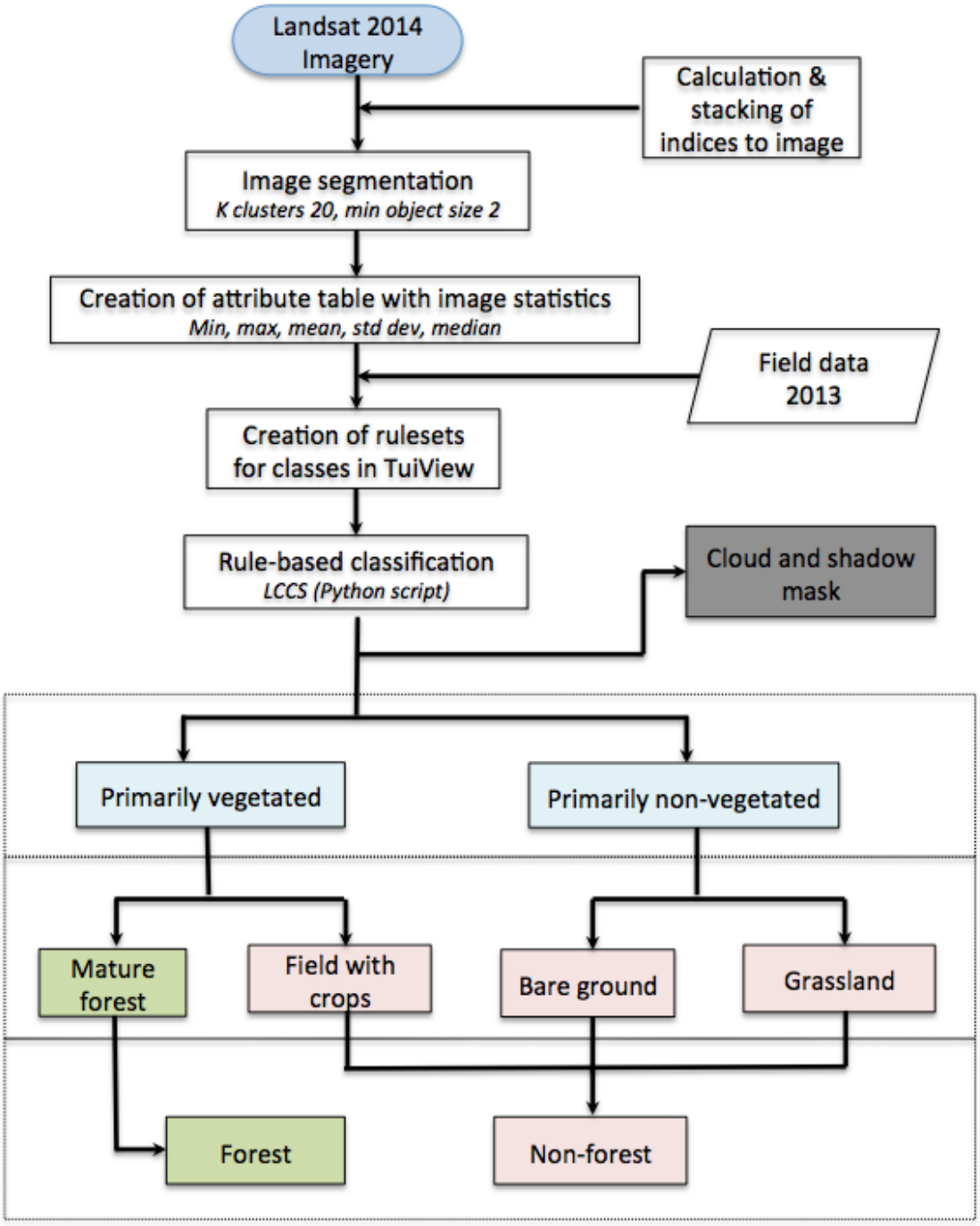


Figure 7.1: The schematic overview of the developed approach for classifying land covers adopted from the rule-based LCCS classification for Landsat data.

LCCS Classification

The EODHaM semi-automated classification system that was adapted from FAO's LCCS was the approach that was used to produce a classification system, following the same steps undertaken when producing a forest baseline from RapidEye imagery. The LCCS and EODHaM systems are flexible and therefore allow for transition to the lower resolution of the Landsat imagery.

Using the statistics in the RAT attribute, unique rulesets were developed for each class as laid out in Chapter 5 and illustrated in Figure 5.6 and Figure 5.7. Unlike in RapidEye imagery where it was noted that the southern part of the Liwale site was slightly different from the northern part so that the ruleset for the forest class had to be slightly adjusted between the two locations in order to efficiently capture the class, one ruleset was found to be adequate for Landsat. Table 7.1 is a summary of the different rulesets (ranges) of the respective spectral bands and NDVI that was used to run an LCCS classification, and thus produce a forest baseline of Liwale.

Table 7.1: The developed rulesets for classifying for the forest baseline of Liwale from Landsat optical data at the different levels of the LCCS.

Level	Class	Ruleset
No data mask	Cloud	Coastal mean >35 Blue mean >60 SWIR1 mean >170
	Shadow	Coastal mean <30.6 SWIR1 mean <120
LCCS Level 1	Primarily non-vegetated	NDVI Mean <100.65
	Primarily vegetated	NDVI mean >100.64
LCCS Level 3	Mature forest	Coastal mean <22.6 Blue mean <22 SWIR1 mean <165 NDVI mean >100.77
	Filed with crops	Blue mean >21.9 & <25.2 NDVI Mean >100.5
	Bare ground	Blue mean >24.9
	Grassland	Coastal mean <6 NDVI mean <100.77

For Landsat, it was found that not as many land classes were identifiable or separable in the landscape. For example, unlike in the RapidEye where there was a clear separation between mature and productive forest in the forest class, these were not separable in the forest class for Landsat. The same was noted in bare fields (without crops) and bare ground/grassland which had were separable and spectrally different in RapidEye, these classes were found to be spectrally similar in

Landsat imagery. On the contrary, even though the separable LULC classes were fewer and more generalised for Landsat compared to RapidEye where a detailed separation and classification could be performed, it was found that simpler rulesets were needed to perform a classification, and these are shown in Table 7.1. Two spectral bands that RapidEye imagery does not have and were found to be effective for separating forest and grassland classes are the coastal and SWIR bands. These rule-sets were then used to run a classification on the RAT image using a python script.

7.3.2 Forest Change Detection

The 2014 forest baseline was used to perform a classification-to-image change detection analysis. As discussed in Chapter 4, change detection for monitoring purposes usually entails going forward in time to see what change has taken place and where (Alqurashi and Kumar, 2013; Mancino et al., 2014). Ideally, an image that had been acquired during the wet season of 2015 would have been used with the 2014 baseline for this analysis. However, in the Landsat time series of Liwale, it was found that due to cloud cover and seasonality, suitable imagery was not available and therefore an image from 1998 (i.e., back in time) was used to demonstrate the change method. Figure 4.11 demonstrates that there was no appropriate image for the 2015 wet season. Landsat has a repeat cycle of 16 days, meaning there are only eight possible images per wet season, and all these were inappropriate.

This, therefore meant that the technique was applied in reverse as change was being tracked back in time. As illustrated by Figure 7.2, for RapidEye data, a 2012 baseline was used with a later date image to detect change. But for Landsat data a back-tracking approach was adopted, meaning change was tracked in reverse. First, a forest baseline was established for the year 2014 using a 2014 image. The baseline was then used, together with a 1998 image (back in time) to produce a change map for the period 1998-2014. This change map was then used to produce a 1998 forest baseline for Liwale. The period between the images used in RapidEye was only two years, thus the detection of forest loss only. However, the Landsat images that were used were acquired 16 years apart. The lengthy years between the two images provided a longer temporal baseline for the change detection, allowing both forest removal and growth to be considered. So. unlike for change detection using RapidEye imagery where only forest to non-forest change was possible, the detection of forest gain (regrowth) was also possible.

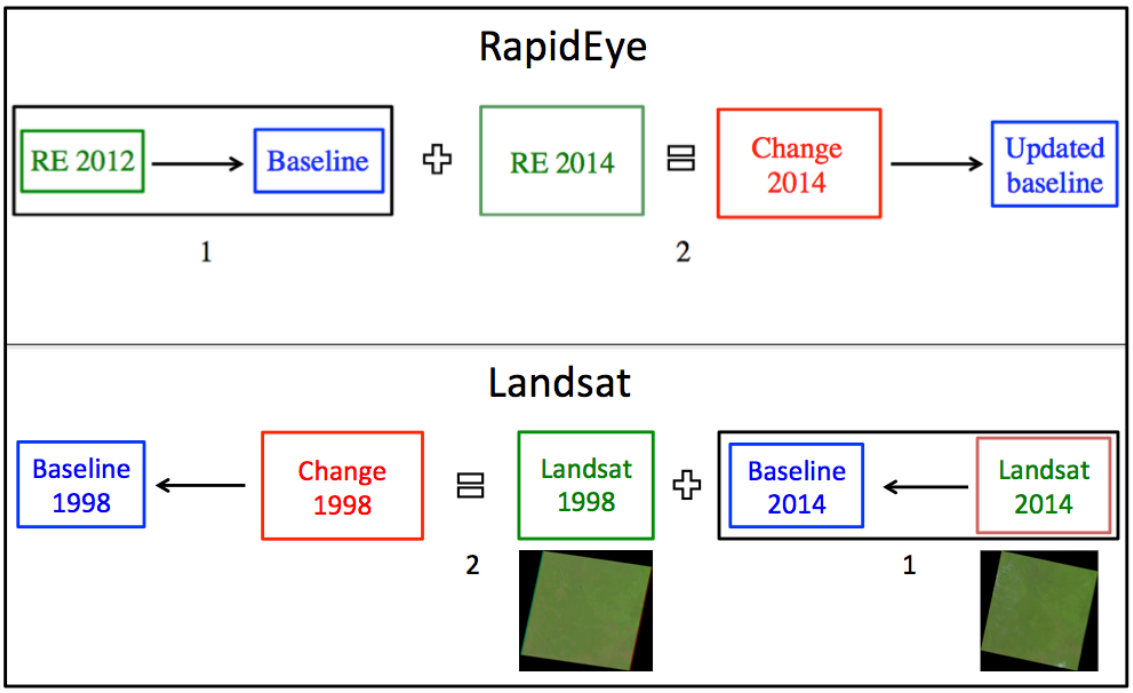


Figure 7.2: Illustration of the back-tracking approach that was used with Landsat data for change detection in Liwale, unlike in RapidEye data where change was tracked forward in time.

Indices and Data Thresholding for Potential Change Features

Unlike in the case of RapidEye where change was only one-fold, the change was two-fold, which meant that there were two sets of potential change features to be identified. First,the 2014 non-forest class data was thresholded to identify potential forest loss features. Secondly, the 2014 forest class data was thresholded to identify potential forest gain features. Figure 7.3 shows the SWIR1 and NDVI histograms, which were the band and index that were used to separate the potential forest loss features from the dataset. These histograms are a plot of all the objects that had been classified as non-forest in the 2014 forest baseline, and the tails circled in red are mainly the objects that were believed to have experienced change over the years in the non-forest class.

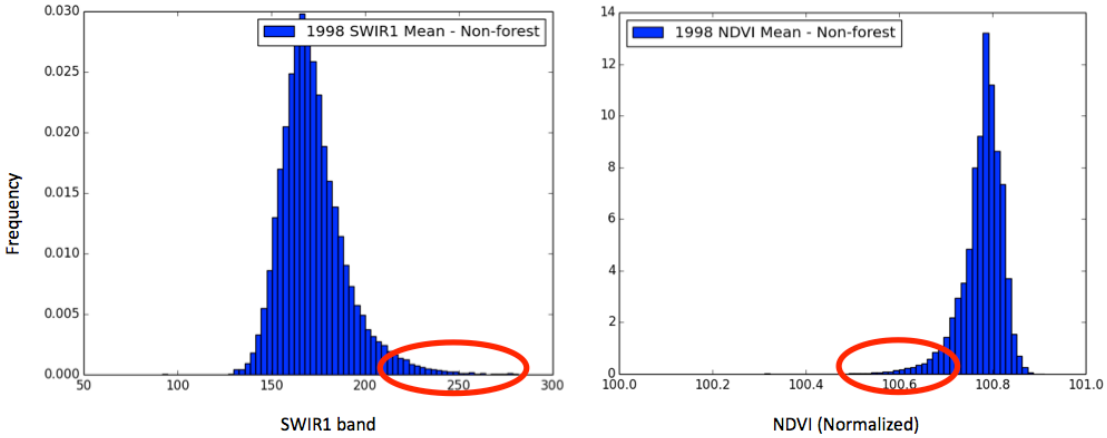


Figure 7.3: SWIR1 and NDVI histograms confirming the tail of potential change features from the normal distribution of the non-forest class.

Similarly, the whole dataset of the objects that had been classified as forest in the 2014 forest baseline was plotted for the spectral bands and NDVI, and it was found that the blue and green spectral bands offered the best separation of the potential change features from the rest of the dataset. Figure 7.4 are the histograms of these spectral bands, and circled in yellow are the tails of the potential change features that had to be thresholded. Then using Skewness and Kurtosis, the potential change features were thresholded for both forest lost and gain.

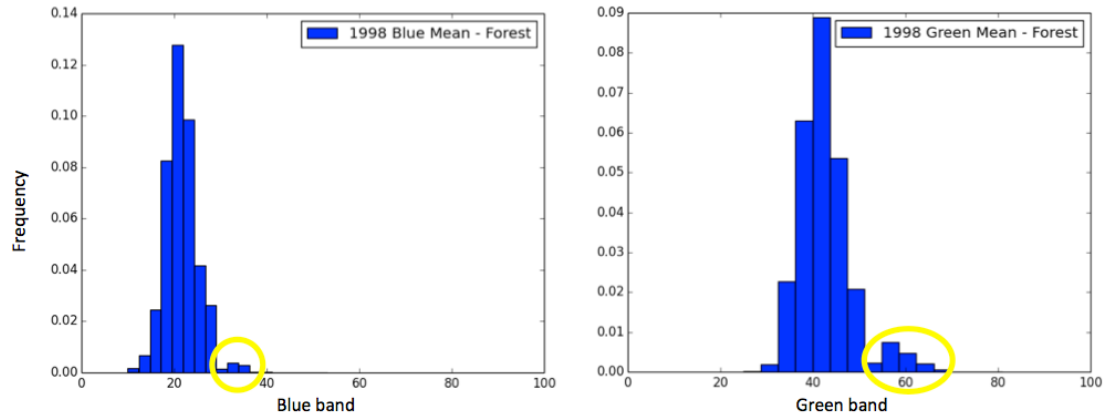


Figure 7.4: SWIR1 and NDVI histograms confirming the tail of potential change features from the normal distribution of the forest class.

Classifying Potential Change Features

After potential change features had been identified, the random forests classifier was used to run a change classification, following the same procedure that had been used in Chapter 6. With one of the aims being to establish how scalable were the methods developed for the high resolution RapidEye imagery to the low resolution Landsat imagery, the same input parameters that had been established for RapidEye were used. These were the input number of trees being 1000, and the gini index was used to determine the number of splitting variables. In scikit-learn (Pedregosa et al., 2011), a classification was performed for the savanna forests of Liwale to produce a change map.

7.4 Results

From the LCCS classification, a forest baseline based on the wet season Landsat image of Liwale was produced, which was then used together with a 1998 wet season image to perform a change classification. Below are the results of the forest baseline for the year 2014 and the forest change that took place between 1998 and 2014.

7.4.1 Forest Baseline Derived from Landsat Image

The western part of the scene was dominated by cloud cover and shadow, and these were masked out and assigned no data value in the classification. The results are presented in Table 7.2 and the numerous figures below.

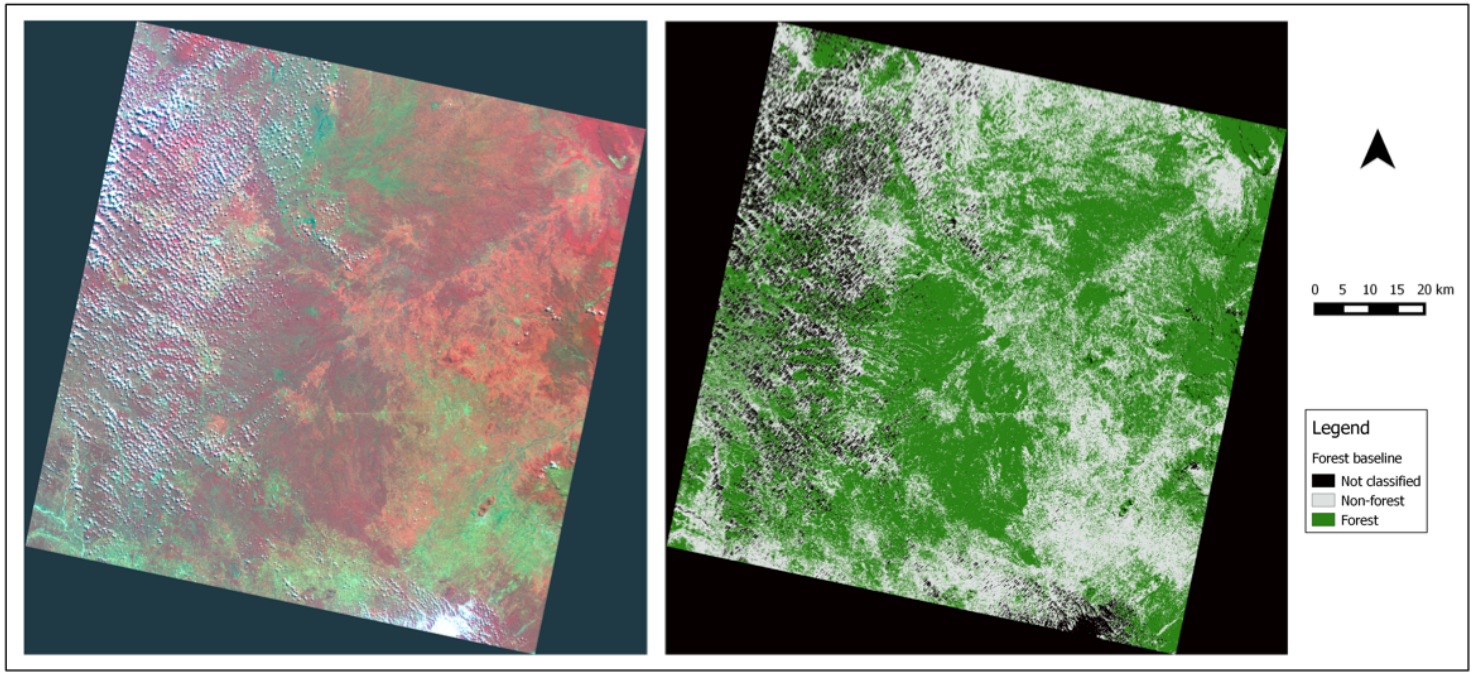


Figure 7.5: The forest baseline of Liwale for the year 2014 produced from Landsat data, and Landsat false colour image displayed as B:NIR, G:SWIR1, R:Blue.

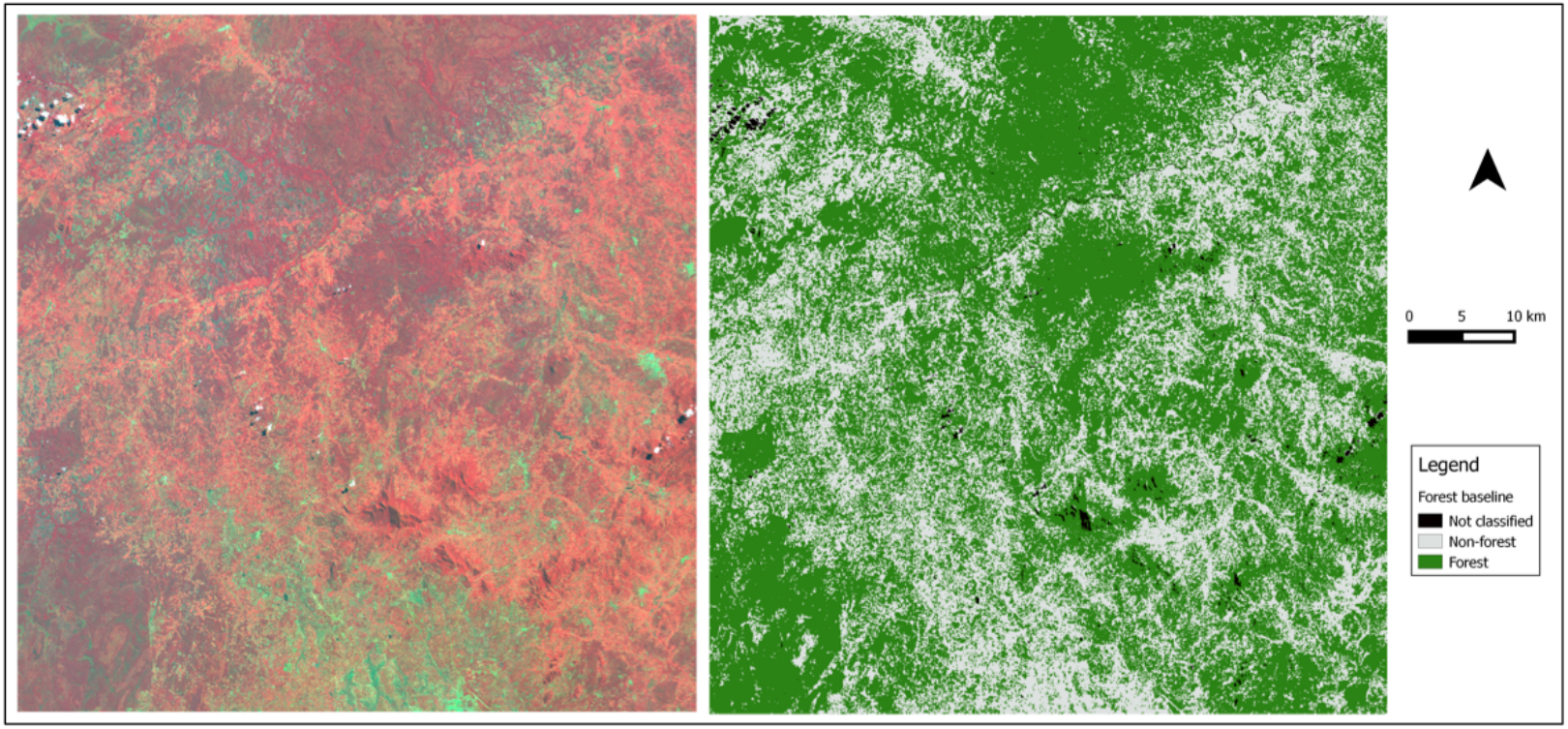


Figure 7.6: A zoom into an area within the Landsat scene that is forested and also experiencing a lot of disturbance from agricultural activities. The Landsat false colour image was displayed as B:NIR, G:SWIR1, R:Blue.

Table 7.2 shows that the forest class covers almost 50% of the scene, it covered 18615 km^2 in 2014. Non-forest covered 15093 km^2, and the remaining 3767.8 km^2. Both the forest and the non-forest classes were irregularly spread across the scene. The non-forest class was dominated by agricultural areas concentrated around the same location, especially to the eastern part of the image. Most of the western part of the scene is protected areas under the Selous Game Reserve, which may be the reason for the less observed human related activity on that part.

Table 7.2: The coverage of the land use and cover classes in Liwale for the year 2014.

Class	**Area (km^2)**
Not classified	3767.8
Non-forest	15093.6
Forest	18615.9
Total	37477.3

Overall, the classification did well and was effective in the separation of the classes being classified for, and thus the production of the forest baseline. Figure 7.7 shows a part of the scene that has both the forest and non-forest classes. Circled in blue is a forest area which is almost spectrally identical to areas that are cultivated and with crop. However, as can be seen across the are, these bright forest areas were correctly classified into the forest class, and the fields were correctly classified into the non-forest class, as seen in the area circled in red. In that part of the image, there is a small portion of forest, with the majority of the area being fields with crop. On the baseline, these two have been well separated and correctly classified into the correct class.

The accuracy of the classification and the detail of the forest baseline is subject to the quality and resolution of the imagery being used. As the initial step is to segment the image into objects, the resultant objects are subject to the image and spectral resolution. The detail to which a classification will capture information is dependent on the resultant objects from the segmentation process. Figure 7.8 shows the RapidEye image and the Landsat image, which immediately shows that RapidEye has got a lot of detail that is immediately lost in Landsat due the image being of a much lower resolution. The resultant baselines respectively show the effects of this difference. It shows that RapidEye provide a baseline with much detail, capable of capturing even individual trees and gaps within the forest block. However, Landsat is capable of effectively capturing big objects, but it is proving not adequate for small features in these savanna environments.

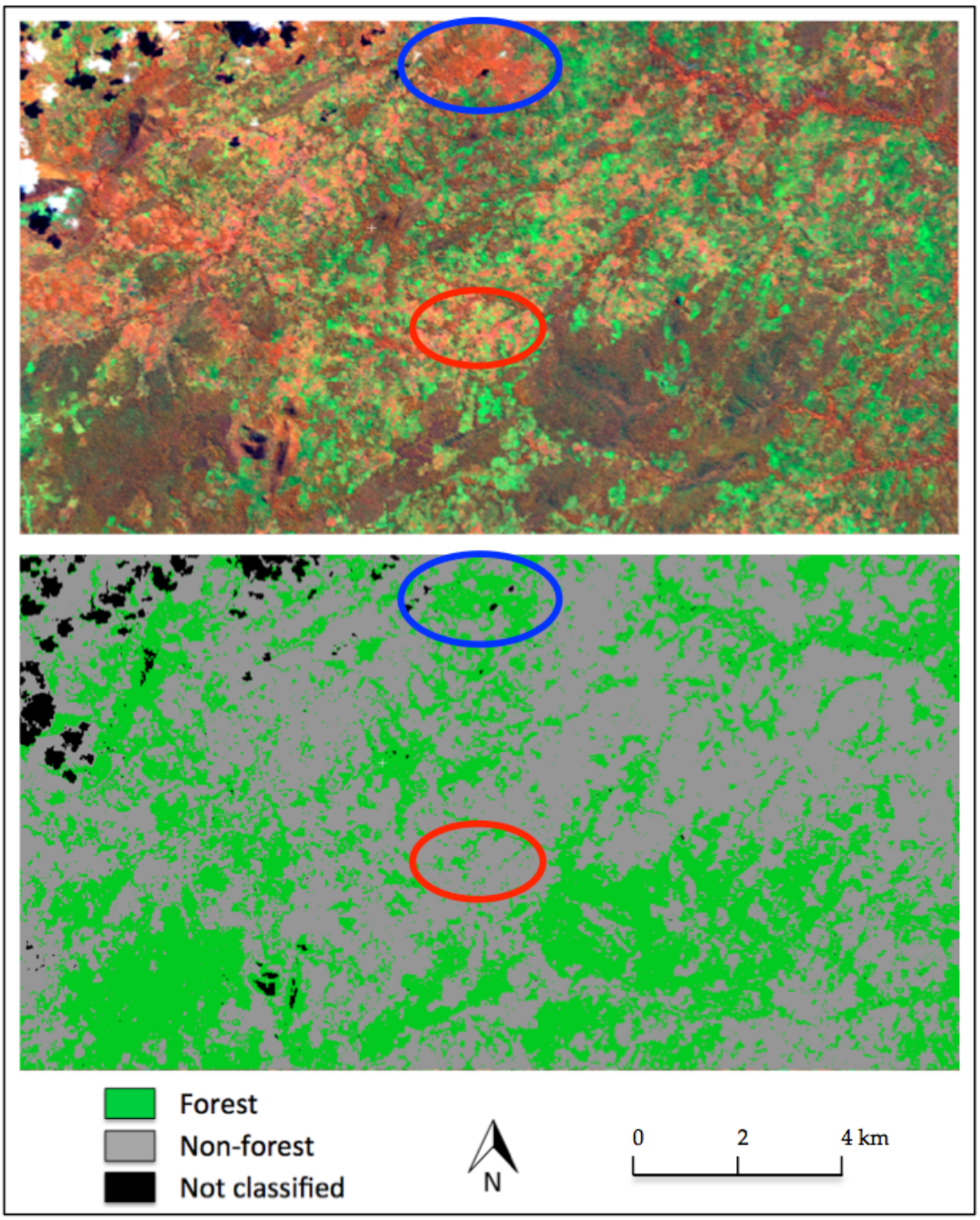


Figure 7.7: A comparison of the forest baseline produced from Landsat and RapidEye images, showing the difference in the detail of the landscape and features the two sensors provide. The Landsat false colour image was displayed as B:NIR, G:SWIR1, R:Blue.

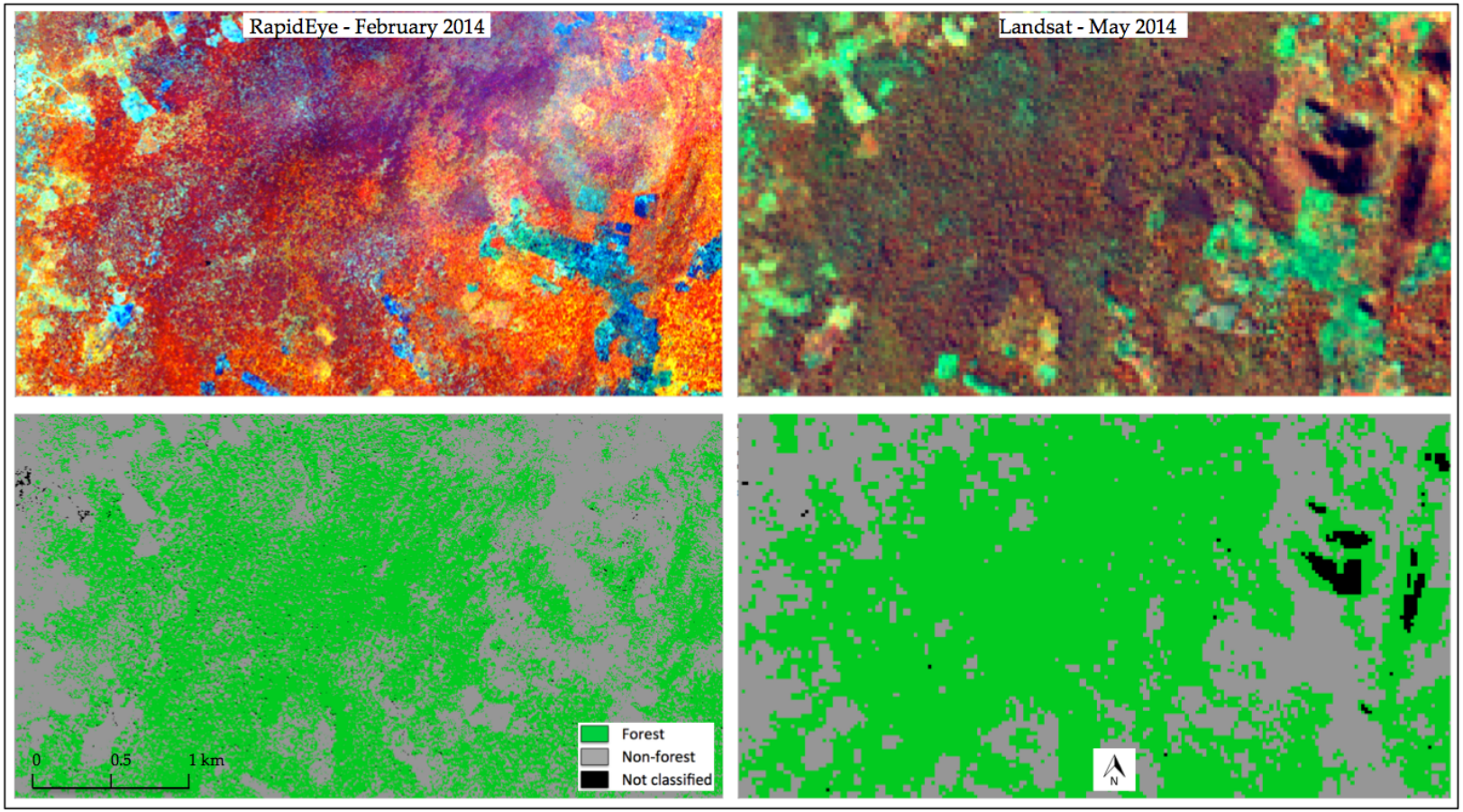


Figure 7.8: An illustration of an area where the forest and non-forest classes have been well classified, even though there is a lot of agricultural activity taking place. The RapidEye false colour image was displayed as B:RedEdge, G:NIR, R:Blue, and the Landsat false colour image was displayed as B:NIR, G:SWIR1, R:Blue.

7.4.2 Validation of Forest Baseline

Using the approach explained in Chapter 5, a sample of validation data was randomly selected across the site. The sampling process involved first generating a shapefile of random points from the produced 2014 forest baseline, using a function in the classification module of RSGISLib called *generateRandomAccuracyPts*. Having assigned the forest baseline as the input image, each generated point was populated with its cover class (forest or non-forest) in the attribute table of the resultant shapefile. Then in QGIS software, an accuracy assessment tool called Class Accuracy (Bunting, 2015) that had been installed as a plug-in in QGIS was used to validate the sampled points. Validation in the tool was done through comparing the assigned class in the classification for each point against the reference data, as had been done in Chapters 5 and 6.

Table 7.3 is a summary of the overall, user's and producer's accuracies for an assessment based on the set of 6000 randomly selected sample points spread across the scene. The overall accuracy was 90.94%, with a kappa coefficient of 0.819. The classification had acceptable user's and producer's accuracies. The classification produced user's accuracies of 93% for the forest class and 91.8% non-forest classes, while the producer's accuracies were 92.3% and 92.6% for the forest and non-forest classes respectively (Table 7.3).

Table 7.3: Error matrix of the overall accuracy assessment for the Landsat-derived forest baseline of Liwale for the wet season of 2014.

Class	***Forest***	***Non-forest***	***Total***	**User's (%)**
Forest	2868	215	3083	93.03
Non-forest	238	2679	2917	91.84
Total	3106	2894	6000	
Producer's (%)	92.34	92.57		**92.45**

To qualify these accuracy measures as recommended by (Olofsson et al., 2013), these were further presented with 95% confidence interval, as explained in Chapters 5 and 6. The forest baseline map had an overall accuracy of 0.92 ± 0.006. It was noted that the user's and producer's accuracies produced using the image-based error matrix were consistent with those presented with 95% confidence. Moreover, uncertainties were quite low, with the highest being ±1%. These accuracies were relatively similar to those derived from the use of an error matrix from the baseline map.User's accuracies were 93% for the forest class and 92% for the non-forest class, with uncertainties of ±0.9% and ±0.8% respectively. Producer's accuracies on the other hand were 90% for the forest class and

94% for the non-forest class, with uncertainties of ±1% and ±0.7% respectively. A summary of the accuracies and their 95% confidence are presented in Figure 7.4.

Table 7.4: Error matrix for the 2012 forest baseline map of Liwale derived from RapidEye data. Map categories are the rows while the reference categories are the columns. The accuracy measures are presented with a 95% confidence interval and the overall accuracy was 0.92 ± 0.006.

Class	*Forest*	*Forest loss*	*Total*	*Am [sq. km]*	*Wh*	*User's*	*Producer's*
Forest	2868	215	3083	13846	0.44	0.93±0.009	0.9±0.010
Forest loss	238	2679	2917	17789	0.56	0.92±0.008	0.94±0.007
Total	3106	2894	6000	31634	1		

7.4.3 Change Map from Landsat

Below are the results of the forest change analysis over Liwale between the years 1998 and 2014, based on the low resolution Landsat data. Figure 7.9 is the 2014 forest baseline and the change map of the same coverage area between the years 1998 and 2014. The map shows that the change areas are concentrated along the main road towards Dar es Salaam. Other features that change concentrate along are rivers and flood plains. These are generally links/features prone to human settlements developing along them for easy access to services, as well for easy access to a water source, both for drinking water and for crop production.

In the year 1998, the forest class covered 19879 km^2, but had reduced to 17788.93 km^2 by the year 2014. This saw a forest loss of 2989.91 km^2. A forest loss meant an increase on the non-forest class, which saw an increase of the class from 11755.14 km^2 in 1998 to 13845.52 km^2 by the year 2014. Regrowth in the area accounted for a forest gain of 899.53 km^2 (Table 7.5).

Table 7.5: The change statistics of the forest extent of Liwale between 1998 and 2014, derived from Landsat data.

Class	**Area (km2)**	
	2014 Forest Baseline	*1998 Change Map*
Not classified	3358.94	3358.94
Non-forest	13845.52	11755.14
Forest	17788.93	19879.00
Forest loss		2989.91
Forest gain		899.53

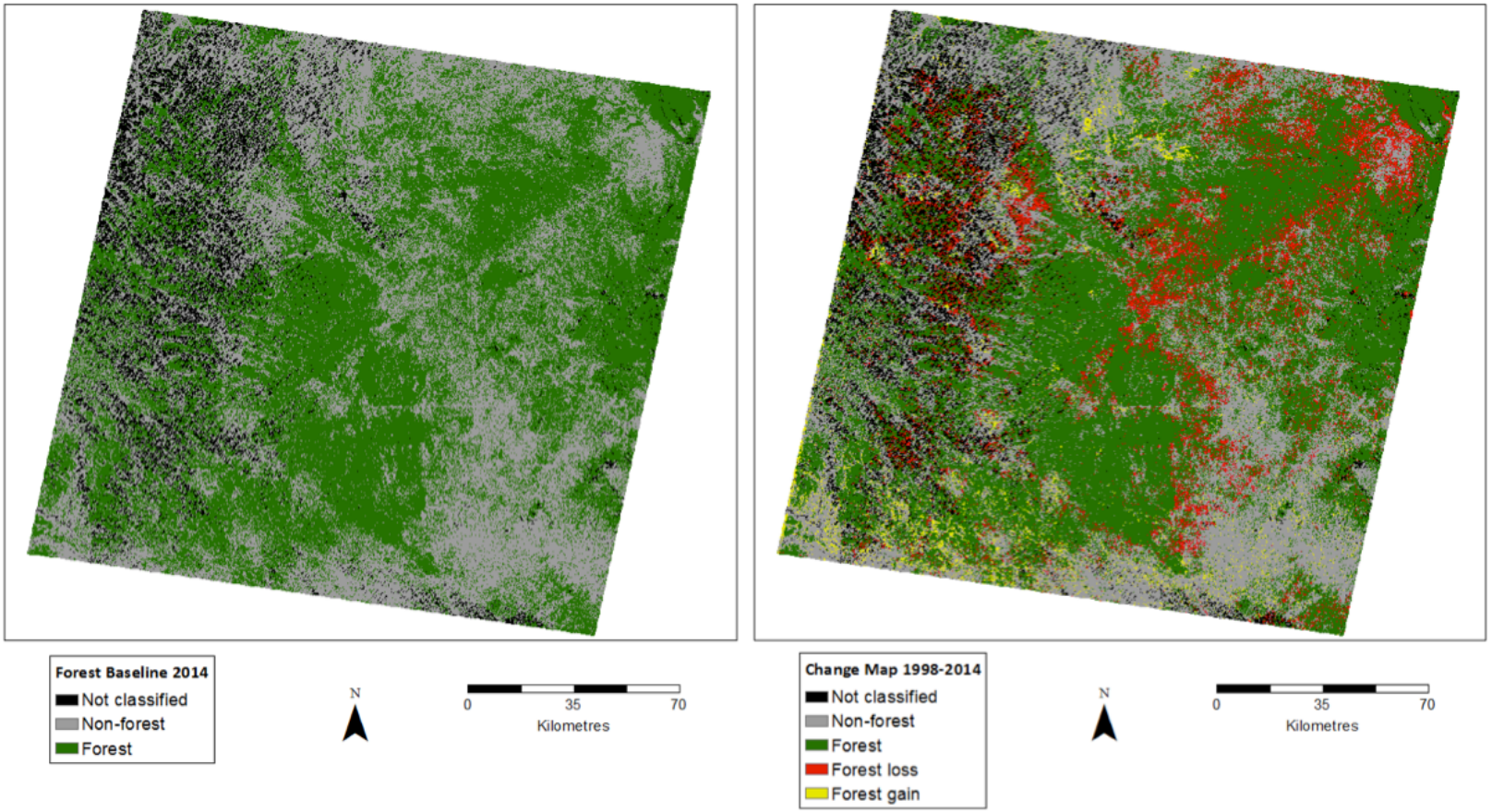


Figure 7.9: The change map of the Landsat scene over Liwale between the years 1998 and 2014, derived from Landsat data.

Figure 7.10 and Figure 7.11 are subsets of the site showing in detail the amount of forest that has been lost between 1998 and 2014. In Figure 7.10, it shows how an area that was dominantly forest in the year 1998 had been extensively cleared by 2014. The amount of forest area that was lost between the years is shown by the red colour in the 1998 change map. Because of the lengthy time between the acquisition times of the two images (about 16 years), regrowth had been experienced in the Liwale site, and in yellow are the regrowth areas.

Figure 7.10 highlight how the environment is being shaped by shifting cultivation that is practised by subsistence farmers in the region and the country at large. The blue polygon shows that in the 1998 false colour image, the area was heavily cleared, probably being the agricultural fields being cultivated then. But by 2014, a bigger portion of the area has been cleared, and the originally cleared area has been left to regrow back to a forest. The classification has correctly captured it as it is classified as forest gain in the change map. On the contrary, the area shown with a yellow polygon was initially forest in 1998 as shown in the false colour image. But by 2014, a large part of it had been cleared for agricultural purposes, and those parts are classified in red on the change map (Figure 7.10).

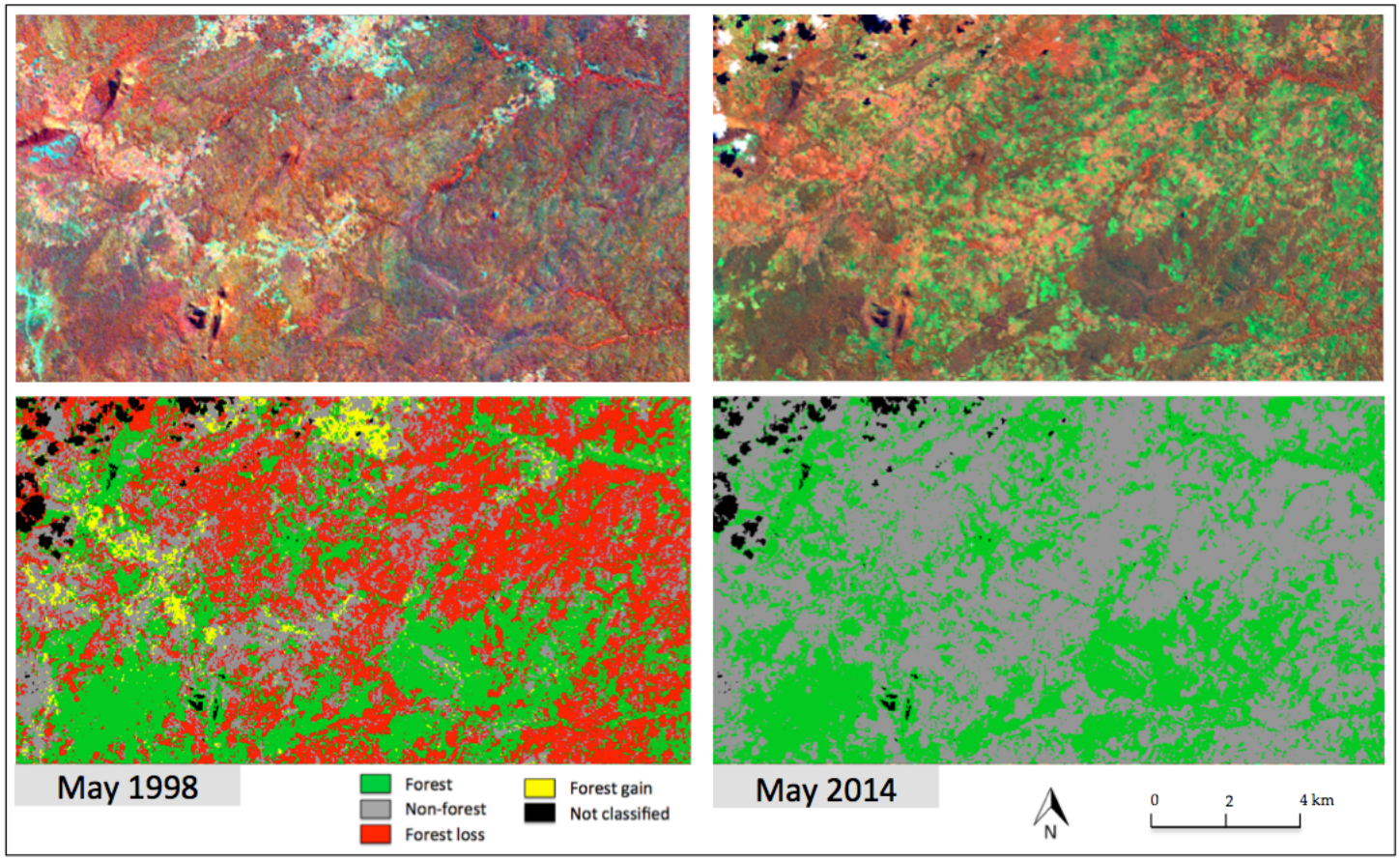


Figure 7.10: A subset of Liwale site showing the high amount of change that has been experienced in some parts of the area. The Landsat false colour image was displayed as B:NIR, G:SWIR1, R:Blue.

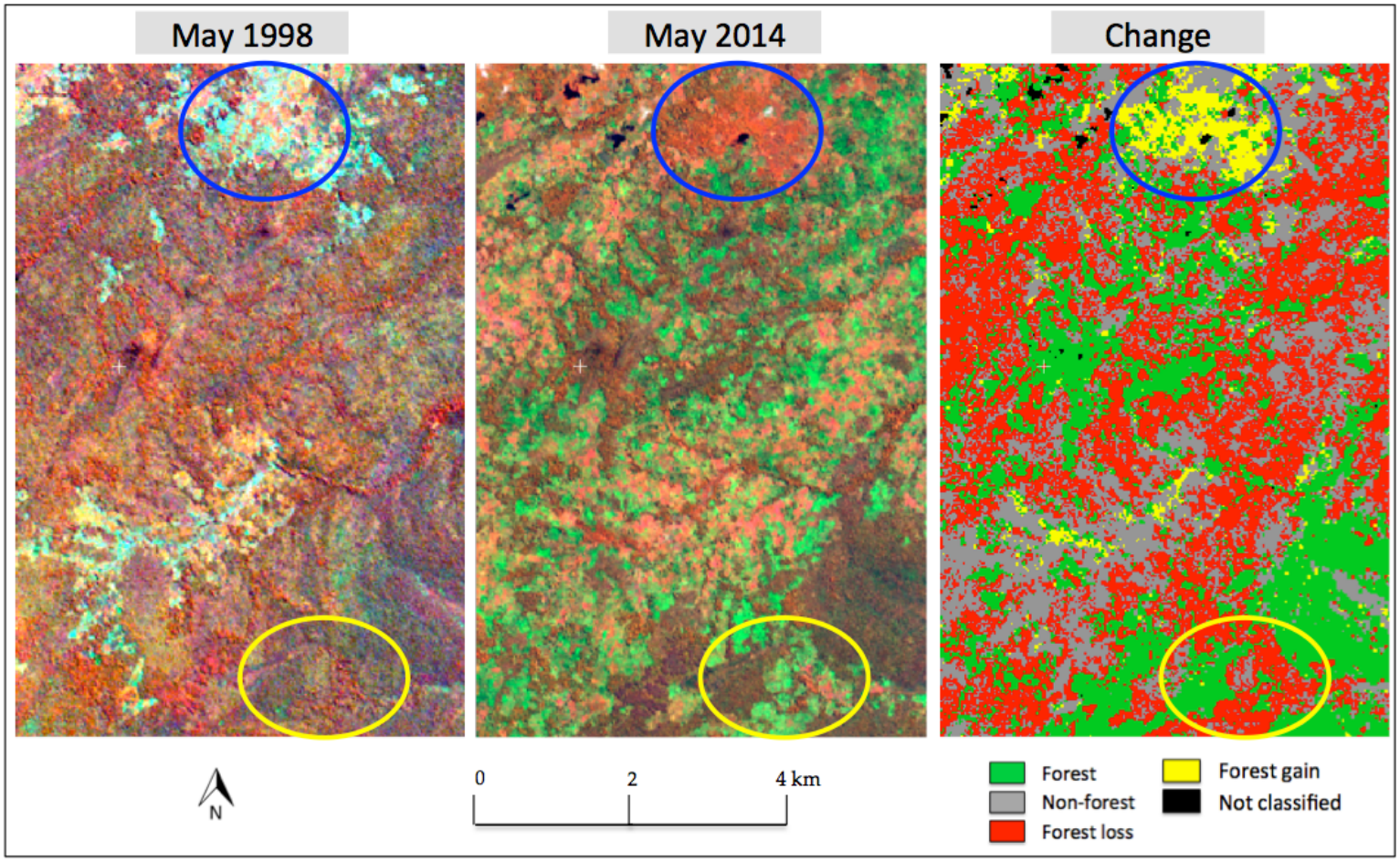


Figure 7.11: The blue circle shows an area that was being cultivated in 1998, but has gone to forest in 2014. The red circle shows an area that was predominantly forest in 1998, but has been cleared extensively in 2014. The Landsat false colour images were displayed as B:NIR, G:SWIR1, R:Blue.

7.4.4 Updated 1998 Forest Baseline

Through adding the gained forest objects into the forest class, and adding the lost forest objects to the non-forest, an 'updated', rather backdated forest baseline of Liwale for the year 1998 was produced. Figure 7.12 are the forest baselines for the years 1998 and 2014 respectively. The forest loss in the north-easterly direction is notable, which as noted in the change map, was primarily due to forest clearing for agricultural purposes along that line of human settlement. Generally, in 1998 the forests of Liwale were relatively closed, but the 2014 baseline shows extensive fragmentation of these forests. This again shows the extent of human interaction and widespread disturbance on the environment that has taken place over the years since 1998.

The observations drawn from the baselines map lines up with the statistical figures of the forest and non-forest classes in the area, as shown in Table 7.6. Even though was regrowth, and thus forest gain in the area, the forest of Liwale was reduced from 19,879 km^2 to 17,788.93 km^2, a decrease of 10.52% in the 16 years. With the forest loss meaning that land was converted to non-forest, this class saw an increase of 17.78% over the years, increasing from 11,755.14 km^2 to a total coverage of 13,845.52 km^2 (Table 7.6).

Table 7.6: Updated 1998 forest baseline of Liwale versus the 2014 forest baseline.

Class	**Area**		
	2014 Forest Baseline(km^2)	*1998 Forest Baseline (km^2)*	*Loss/Gain (%)*
Not classified	3358.94	3358.94	
Non-forest	13845.52	11755.14	+ 17.78
Forest	17788.93	19879.31	- 10.52
Total	34993.39	34993.39	

Figure 7.13 is a zoom into an area that has seen high clearing over the years. The forest baseline of the area shows that in 1998, the area was mainly forested, with over 75% of it covered in forest. However, by the year 2014, there had been increased human settlement and clearing of the forest. The overall forest cover had reduced to less than 40%. The forest cover percentage could have been even further, however, with shifting cultivation being popular among subsistence farmers, some abandoned areas had regrown to improve the forest cover in this area. An area of forest regrowth is shown with a blue polygon, which shows that it was cleared in 1998, and correctly classified as non-forest in the 1998 baseline. However, the polygon area had reverted to a forest area by 2014, as seen on the false colour image, and this was correctly classified as forest 2014.

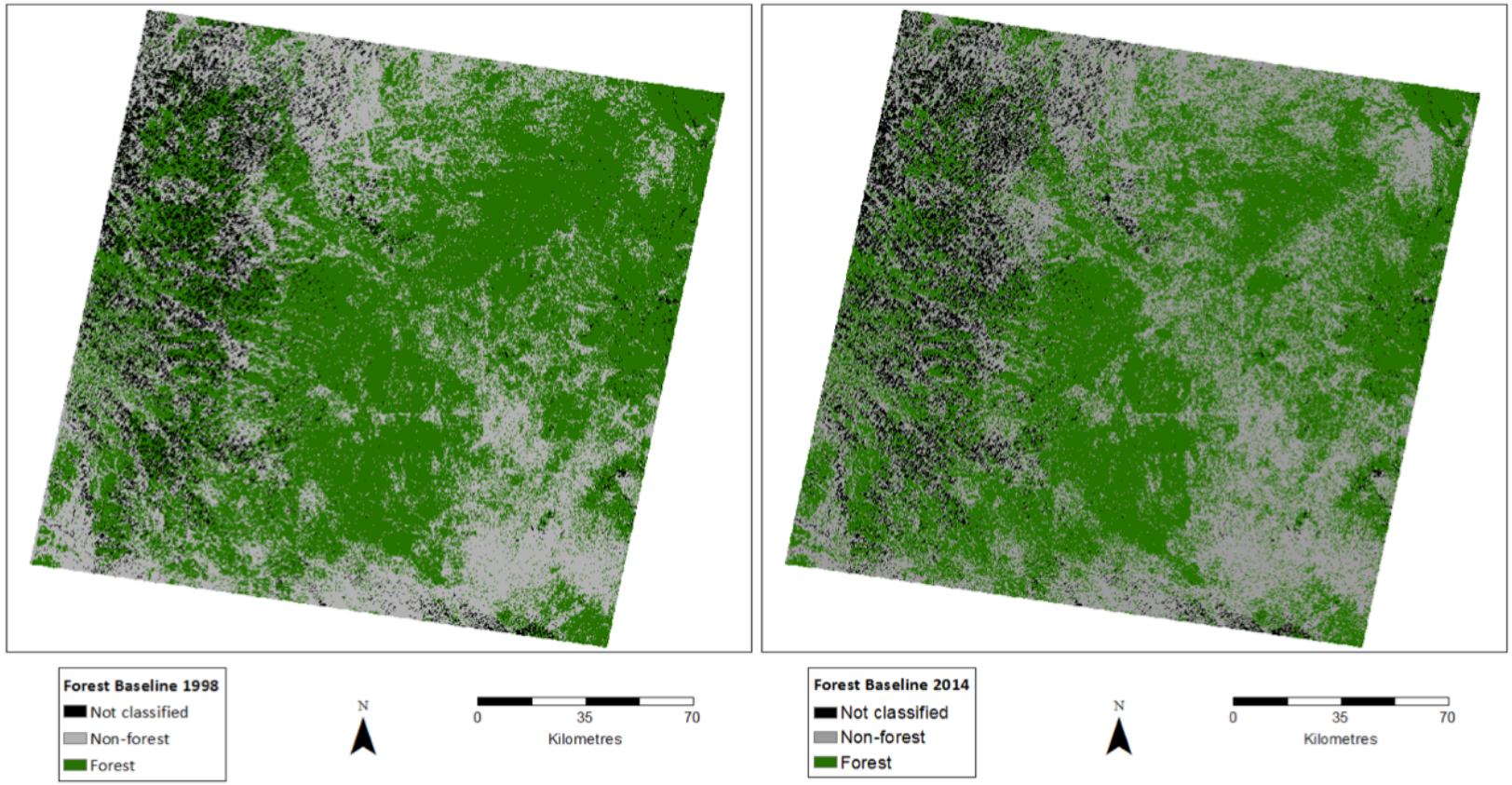


Figure 7.12: Forest baselines for the years 1998 and 2014 over Liwale.

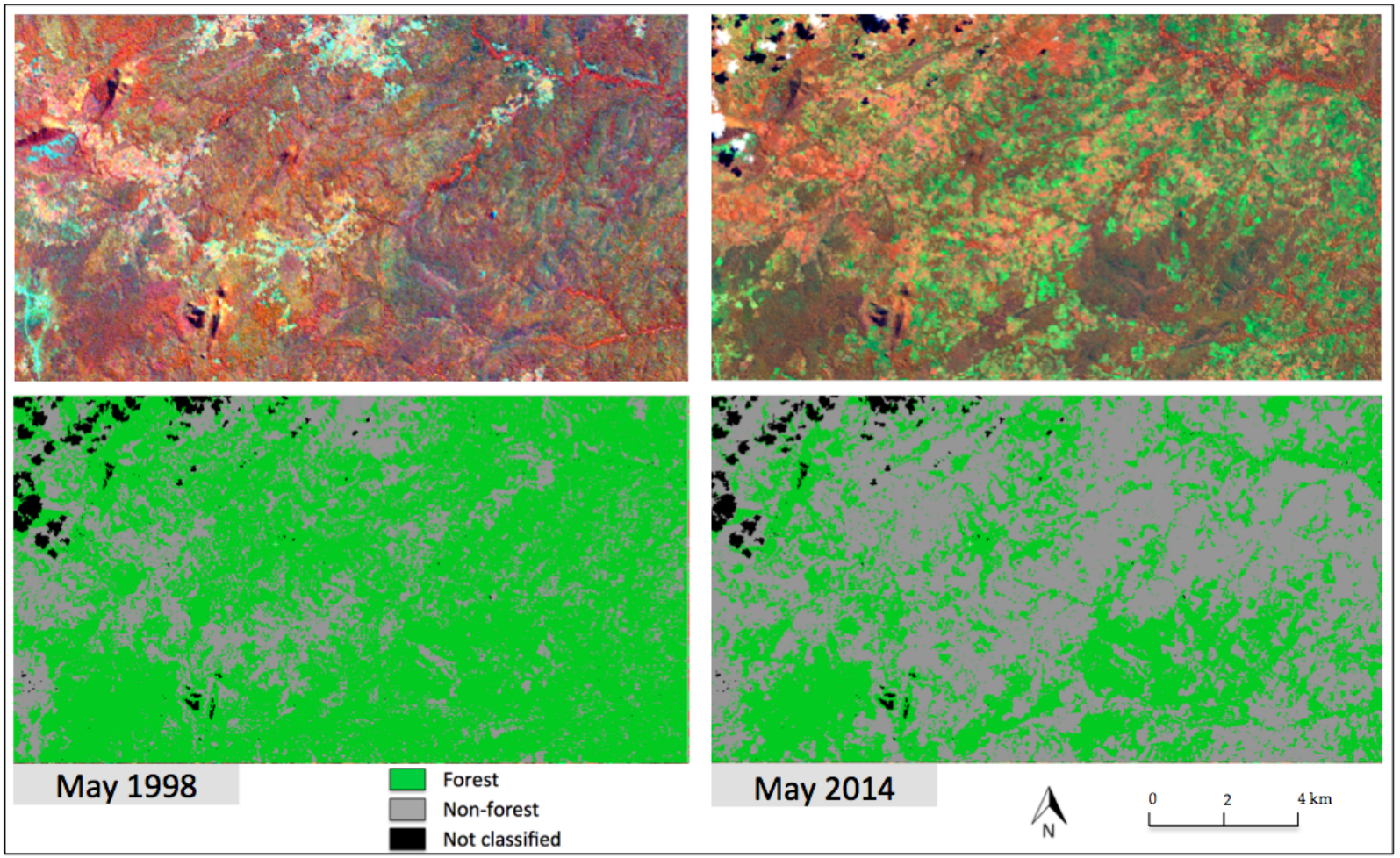


Figure 7.13: An updated forest baseline for the year 1998 against that of the year 2014 for a subset of the Liwale site. The Landsat false colour images were displayed as B:NIR, G:SWIR1, R:Blue.

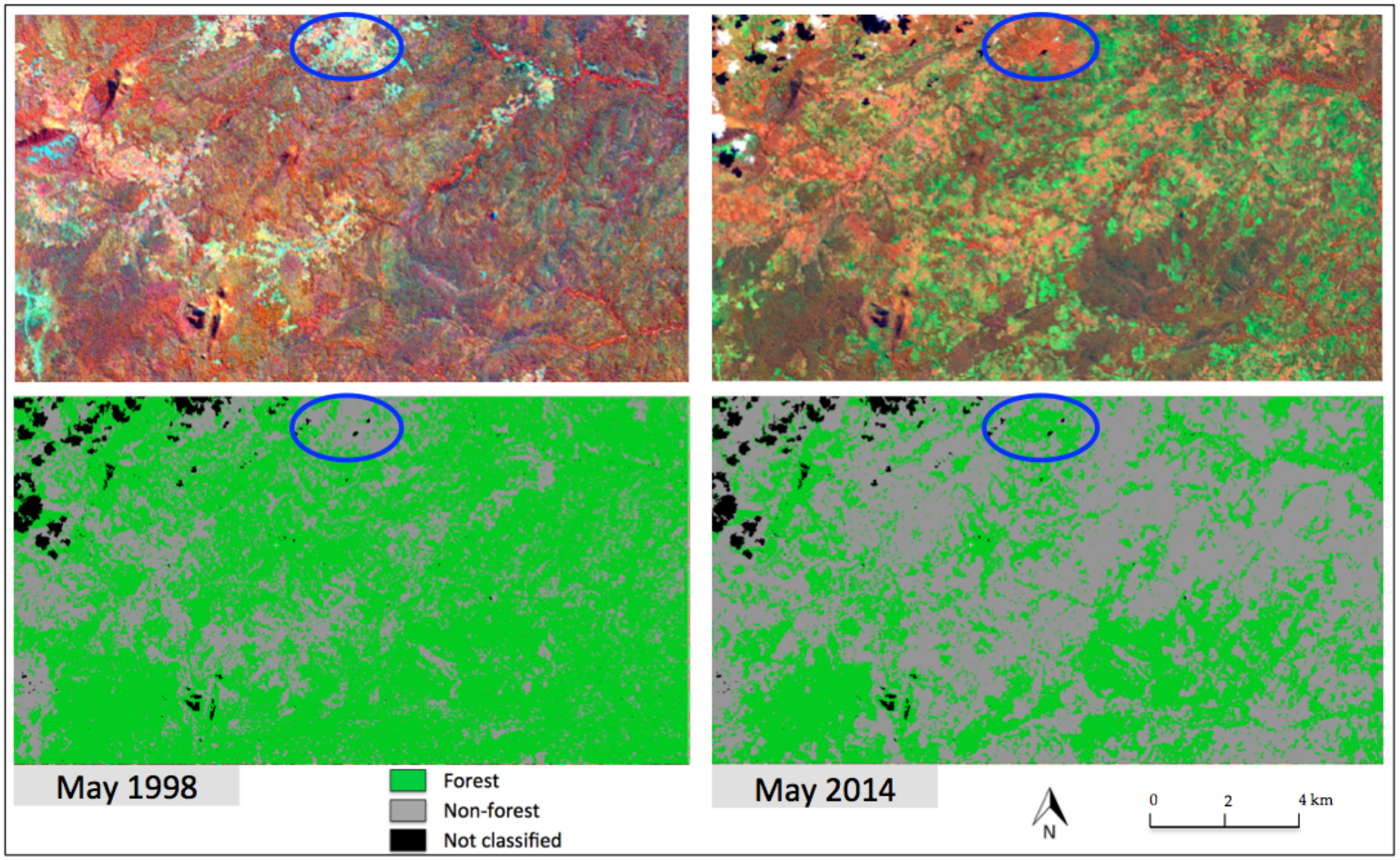


Figure 7.14: An area (circled in blue) that was being cultivated in 1998 and thus classified as non-forest in the baseline, but has gone to forest in 2014 and thus classified as forest in the 2014 baseline. The Landsat false colour images were displayed as B:NIR, G:SWIR1, R:Blue.

7.4.5 Validation of Change Results

To assess the accuracy of the change result derived from Landsat images, a stratified random sampling approach was adopted. Unlike in RapidEye where possible change was only moving from forest to non-forest, or a forest area experiencing disturbance in the form of degradation, due to the lengthy time between the baseline year and the image used for detecting change, validation had to be done for forest gain areas.

First, two binary images were created, one for the 2014 forest class, and another for the 2014 non-forest class. The 2014 forest binary was used to clip all the objects in the 1998 change map that were forest in 2014. The clipped image was then used to perform stratified random sampling for both the forest class and the forest gain class, and 1500 forest random points and 400 forest gain random points were generated. With forest gain being very low in coverage compared to the forest class, the sampling was done to ensure representation of both classes. Similarly, the non-forest binary was used to clip out all objects in the 1998 change map that were non-forest in 2014. Stratified random sampling was performed on the clipped image to produce 1000 non-forest points and 300 forest loss random points.

The set of points were then merged, and a 30 m buffer was created around them. The Class Accuracy plug-in was used to validate each of the sampled points within a window of four Landsat pixels. The validation was done using the 1998 false colour composite image and a Google Earth image. Limitations noted for Google Earth coverage of Liwale, as discussed earlier, necessitated the use of the Landsat composite as the primary validation image, and the Google Earth image as a secondary validation image.

The overall accuracy of the change map of Liwale based on Landsat data was found to be 93.16%, with a kappa coefficient of 0.900, for a total number of 3200 stratified random points. Table 7.7 is the error matrix for the accuracy assessment for the change map. It shows that both the forest gain and loss classes were well detected. The user and producer accuracies for forest loss was 85.3% and 100% respectively, while the user and producer accuracies for the forest gain was 86.5% and 86.9%, still an acceptable level of accuracy.

Table 7.7: Accuracy assessment of the change detection results for the Landsat scene over Liwale.

Class	*Forest*	*Non-forest*	*Forest loss*	*Forest gain*	Total	User's (%)
Forest	38	2387	3	36	2464	96.88
Non-forest	2795	78	0	8	2881	97.01
Forest loss	37	21	384	0	442	86.88
Forest gain	0	44	0	169	213	79.34
Total	2870	2530	387	213	6000	
Producer's (%)	97.39	94.35	99.22	79.34		**95.58**

As recommended by Olofsson et al. (2013), the accuracy measures were further presented with 95% confidence interval, which gives more information on uncertainty. Table 7.8 presents the user's and producer's accuracies for the change map, with 95% confidence interval. The overall overall accuracy of the change map was 0.96 ± 0.005. The change map had user's accuracies of 97% for the forest and non-forest classes, respectively. These two classes also had high producer's accuracies of 93% and 98% respectively. In terms of uncertainty, both classes had uncertainties of a percent (1%) or less. However, the user's accuracies for forest loss and forest gain were lower, with relatively higher uncertainties. These classes had accuracies of 87% and 79%, with uncertainties of ±3.2% and ±5.4% respectively. for the forest loss class, with uncertainties of ±0.5% and ±0.6% respectively. The producer's accuracy for forest gain was the lowest (76%), with ±6.8% uncertainty. On the contrary, forest loss had 99% producer's accuracy, with an uncertainty less than 1%. The result shows that there was actually more omissions on the change map for the forest gain class, which the error matrix could not show without undertaking further statistical analysis.

Table 7.8: Error matrix for the 2012 forest baseline map of Liwale derived from RapidEye data. Map categories are the rows while the reference categories are the columns. The accuracy measures are presented with a 95% confidence interval and the overall accuracy was 0.96 ± 0.005.

Class	*Non-forest*	*Forest*	*Forest loss*	*Forest gain*	Total	Am [sq. km]	Wh	User's	Producer's
Non-forest	2387	38	3	36	2464	11755.14	0.331	0.97±0.007	0.93±0.011
Forest	78	2795	0	8	2881	19879	0.560	0.97±0.006	0.98±0.005
Forest loss	21	37	384	0	442	2989.91	0.084	0.87±0.032	0.99±0.006
Forest gain	44	0	0	169	213	899.53	0.025	0.79±0.054	0.76±0.068
Total	2530	2870	387	213	6000	35523.58	1		

From the change map, the estimated forest loss between 1998 and 2014 in Liwale was 2989.91 km^2, while forest gain was estimated to be 899.53 km^2. Using the error-adjusted estimator (Olofsson et al., 2013), the stratified estimator of forest loss was found to be 2611.88 km^2, while that of forest gain was 940.66 km^2. Table 7.9 is a summary of the areal estimates from the change map

and those area proportions adjusted for error for the different classes. The table shows that there was about 4% omission in the classification of forest gain, while there was about 15% commission in the classification of forest loss.

Table 7.9: The total estimated areal coverage for each class in the change map produced from Landsat data, against the values derived from the stratified error-adjusted estimator.

Class	Change map estimate (sq. km)	Stratified estimator (sq. km)	Difference (sq. km)	Omission/ Commission (%)
Non-forest	11755.14	12253.87	498.73	4.07
Forest	19879	19717.17	-161.83	0.82
Forest loss	2989.91	2611.88	-378.03	14.47
Forest gain	899.53	940.66	41.13	4.37

The standard error of the adjusted area estimate for forest loss was found to be 52.09 km^2. Therefore, the final forest loss area estimate with a margin of error (at approximate 95% confidence interval) was 2611.88 $\pm$ 104.18 km^2. Taking this uncertainty into account, the true area of forest loss could be as low as 2507.70 km^2 or as high as 2716.06 km^2, with 95% level of confidence. For forest gain, on the other hand, the standard error of the adjusted area estimate was 26.07 km^2, giving the final forest gain area estimate with a margin of error (at approximate 95% confidence interval) as 940.66 $\pm$ 52.13 km^2. When this uncertainty is taken into account, the true area of forest gain could be as low as 888.53 km^2 or as high as 992.79 km^2, with 95% level of confidence.

Table 7.10: The final land area estimate with a margin of error (at approximate 95% confidence interval) for the different land classes for the change map produced from Landsat data.

Class	Standard error	St. error of adjusted area estimate (sq. km)	Final area estimate (sq. km)
Non-forest	0.00311	110.42	12253.87 $\pm$ 220.84 km^2
Forest	0.00325	115.59	19717.17 $\pm$ 231.18 km^2
Forest loss	0.00147	52.09	2611.88 $\pm$ 104.18 km^2
Forest gain	0.00073	26.07	940.66 $\pm$ 52.13 km^2

7.5 Discussion

7.5.1 Forest Baseline Classification

For Landsat, it was found that not as many land classes were identifiable or separable in the landscape. For example, unlike in the RapidEye where there was a clear separation between

mature and productive forest in the forest class, these were not separable in the forest class for Landsat. The same was noted in bare fields (without crops) and bare ground/grassland which had were separable and spectrally different in RapidEye, these classes were found to be spectrally similar in Landsat imagery. Two likely causes of these include the lower resolution of the Landsat, and that these two sensors capture data in different spectral bands. RapidEye captures data in the red edge region of the spectrum, and capturing data in this region has been reported to be very effective in the separation of different vegetation types and avails an opportunity to produce indices, and Landsat does not capture data in it. However, the Sentinel 2 instrument will have a high resolution (10 m) and a red edge band and is therefore expected to provide an ideal datasets for this monitoring.

With Landsat, at 30 m resolution, the limitation of heavily fragmented landscape is eliminated, which may be the reason why both the forest baseline and change map produced using Landsat data produced high overall accuracies and kappa coefficients. Even though the classes separable were fewer and more generalised for Landsat, it was found that separation of these classes was very effective, with fewer bands used to build the rule-sets. SWIR was found to be very useful, forming the rule-set for separating most classes. Though a lot of detail on the landscape was lost at 30 m resolution, features were found to be relatively easier to separate for Landsat than for RapidEye.

However, this is a landscape with a lot of subsistence agricultural activities taking place, and sometimes at a very fine local scale. For example, a local farmer may shift from a field into a nearby forested area, partially clearing the canopy over the area. At the Landsat scale, such change is unlikely to be picked. Moreover, depending on the scale of disturbance, Landsat is unlikely to be able to pick up degradation from logging. And the Tanzanian environment being heavily reliant on the natural environment for energy, especially wood-fuel, as well as wood for building material, selective logging has been reported to be rife in some areas. Landsat, as shown in the change detection subsection of Chapter 7, does not provide enough detail to effectively map small scale change features.

7.5.2 Change Detection

Landsat does not have the same ability as RapidEye to give detail of the environment, and thus some detail that would have led to improved estimation of change is lost in the process. However,

at a broader scale, due to the complexity of the environment, the heavy fragmentation creates a feature space that has a lot of overlap. For RapidEye, which at 5 m does not provide adequate information, for example to show the boundary of the canopy of an individual tree, this leads to inadequate segments, and thus a baseline that may have boundary errors. Ultimately, it introduces error into the change map as boundary pixels may be classified incorrectly. However, an overall accuracy of over 93% is very acceptable and shows that the technique holds a great potential for forest monitoring in Tanzanian and African savannas even with the use of low resolution Landsat data. Being a freely available dataset means that a monitoring system build around it would be cost-effective as well.

The issue of poor registration in RapidEye that also introduced an error to both the baseline and change map of the area was not found for Landsat. Both images that were being used for the classification and change detection were noted to be well registered. During the validation process, insights drawn from the Google Earth image included that even though fewer classes in the Modular Hierarchical Phase, the same noted confusion between land classes that had been noted for RapidEye data was found to still hold for Landsat. For example, wetland areas were still found to have been confused with the forest class. Similarly, some forest areas were confused with fields with crop.

Another common cause for confusion within the site was due to intercropping that is commonly practised by subsistence farmers in the area. Most farmers cultivate their crops such as maize and legumes within fields that have cashew trees. Moreover, as highlighted how farmers would rather keep the trees with big trunks in an upright position than fell them, not all the trees are dried up. Moreover, some trees being dried up may take a couple of years before it loses its canopy. These, as illustrated in Figure 7.15, may be hard to separate within an optical image from a forest. In this example, a sampled point was found to lie mainly within a field dominated by trees. Optical data relies on canopy cover to separate classes, especially forest from non-forest, which led to such an area being incorrectly classified as forest.

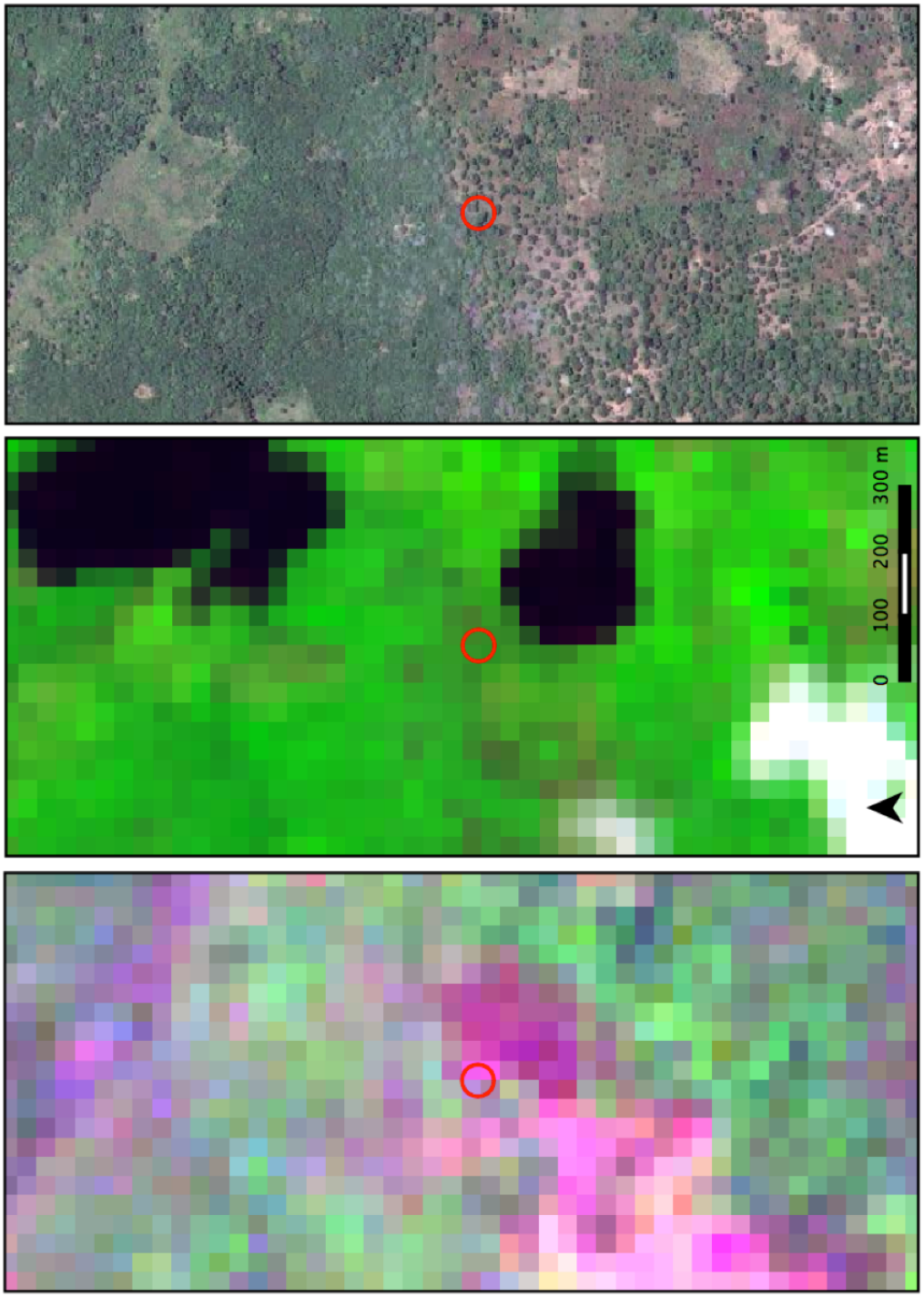


Figure 7.15: An illustration of a sampled point (red circle) falling on the boundary between a field and a forest.

Numerous studies illustrated the usefulness of ancillary data for improving the overall results for both LULC classification and change detection (Rapinel et al., 2014; Drgu et al., 2014; Lucas et al., 2015). As ground data was lacking in this environment, possible options that could be considered include texture and shape. Being mainly a rural landscape, experiencing substantial subsistence agricultural practices, this causes the environment to be irregularly fragmented, without any shapes that can be used to define any features and land classes. As the segmentations show, there are no notable shapes or patterns. This limitations limit the other input that could have been used to improve the results.

Overall, the method was found to be very effective in picking up areas of change, both forest loss and forest gain. It is argued that the method itself works very well for this task, but was limited by data availability and other limitations associated with the environment. Against these limitations and confusions between classes, the overall accuracy of 93.16% and a kappa coefficient of 0.9 shows that the method holds a high prospect of establishing a forest baseline and a monitoring system for these savannas of Tanzania, using the freely available Landsat archive. As mentioned that Landsat does not have the same capability as RapidEye to pick fine detail of loss, such as loss of small features (e.g. a tree or a couple of them), caution need to be exercised when using Landsat. Change detection around and need settled areas, with communities may not be adequate with just Landsat. Big clearings for agricultural practices will be adequately picked and mapped. However, for a tree of just a few pixels lost to, for example, selective logging, forest fires, or animal browsers, these may not be picked up since a bigger part of that pixel may still have enough canopy cover, and thus mapped as non-change. For a national monitoring programme that takes this direction, it may therefore be ideal to not use only Landsat as the primary sensor, but a combination with RapidEye may yield a better and cost-effective result. They could be used to compliment each other, with RapidEye used in areas where there is human activity, and thus higher chances of even small-scale change, and Landsat used in areas where there is not much human interaction with forest areas and change not likely to take place.

In the case of change detection, cloud contamination could lead to false change detection if the affected areas are not effectively masked out. It should also be acknowledged that where extensive areas of cloud are present, the actual area of change becomes unknown, or the certainty of change in those parts of the study area is relatively lower. Fisher (2014) noted that commission errors due to false clouds (e.g. highly reflective ground), and false shadows (e.g. dark water bodies) can be high, as can omission errors due to thin cloud that is very similar to the underlying ground surface

(e.g. haze). Worth noting is that the study further recommended manually editing for such errors, as it was found to be very effective in quickly reducing them. The approach used in the thesis was similar as the masking was done manually, through the use of rulesets, which ensured that the mask was manually validated before a classification was performed.

7.5.3 Flexibility of the Developed Techniques

The successful trialling out of the developed methods for mapping the forest cover and for detecting change in an area from high resolution to course low resolution data shows the flexibility and transferability of these techniques. Even though there are limitations with the use of courser resolution data for such a heterogeneous and sometimes sparsely populated forests, overall, it was found the accuracy of the results were comparable (over 90%). As highlighted by Rapinel et al. (2014) and Drgu et al. (2014), most developed techniques lacked flexibility, and are not transferable between different environments (Wieland and Pittore, 2014). Moreover, the developed techniques tend to be scene-specific or sensor-specific. The results from scaling out these techniques show that these developed techniques are flexible and can be used with different remote sensing data with different spectral bands and resolution.

7.5.4 Remote Sensing Data for Monitoring the Savanna Forests

With both sensors (RapidEye and Landsat) having demonstrated their potential for both forest mapping and change detection (monitoring), it may be a better option to use both of them for large-scale monitoring. There is a high cost associated with the acquisition of RapidEye data, which may limit its use at regional/national scale. In addition, as found for Liwale, RapidEye time series data is very shallow over the country, only dating back to the year 2010. Because of persistent cloud cover, even the relatively small site of Liwale was found to have a very poor coverage by appropriate scenes. Moreover, due to seasonality in the region, all dry season images are not appropriate for both forest mapping and change detection as vegetation cover is poor, being the leaf-off season. Unaccounted for seasonality has the potential to lead one to incorrectly conclude that an object has experienced change. Even within the leaf-on season, the window for appropriate imagery is relatively narrow, ranging between late February to late May. This too is subject to the timing of the rains which is crucial for vegetation to achieve full cover, and when they go for vegetation to start losing their canopy. Cloud cover can be so persistent even during

the wet season where as illustrated in Figure 4.11, there was no appropriate Landsat image for monitoring purposes during the wet season of 2015.

It may therefore be a cost-effective approach to use both sensors interchangeably for forest monitoring in Liwale and greater environments where the proposed system is introduced. This approach would have the ability to reduce the cost of sourcing RapidEye for the whole country, but also ensure that small change is detected effectively using RapidEye. Generally, degradation is rife and likely to happen around settled areas, and within a buffer from them. However, vast forest areas without a lot of human interaction and protected areas are not prone to small scale disturbance. In this case, it is proposed that Landsat be the sensor used for such environments. It has been shown that for larger scale forest mapping and change detection, the sensor is very effective. Then RapidEye data would be sourced for those areas prone or likely to experience small-scale disturbance.

7.6 Conclusions

The methods for mapping the forest baseline of the savanna woodlands, and then using the derived baseline together with an image of a different acquisition date to detect forest change are well suited and scalable to the low resolution Landsat image for the savanna environments of Liwale. It produced both an acceptable forest baseline (with an overall accuracy of 92.45% and a kappa coefficient of 0.819) and a change map of the area (with an overall accuracy of 95.58% and a kappa coefficient of 0.900). A further statistical analysis used to validate the statistical measures showed that the baseline map had an overall accuracy of 92% , with an uncertainty of $\pm 0.6\%$, while the change map had an overall accuracy of 96%, with an uncertainty of $\pm 0.5\%$. Moreover, the margin of error for area estimates was computed, and was found to be 2611.88 ± 104.18 km^2 for the forest loss class and 940.66 ± 52.13 km^2 for the forest gain class.

For the forest baseline mapping, it was noted that the rule-sets needed to separate the different LULC classes were relatively simpler than those required for the separation of classes in RapidEye. These rule-sets were mainly developed from the spectral bands, and only the NDVI was used, unlike in RapidEye where numerous indices were used. However, even though the method is scalable and a good baseline was produced with an a good overall accuracy 93.16%, the resolution of Landsat limits the amount of detail that can be mapped in these heavily fragmented environments. Landsat is

therefore as capable for the establishment of the forest baseline for REDD monitoring in Tanzania, even though it may not account for forest degradation due to lack of detail RapidEye benefits from.

Chapter 8

Discussion

8.1 Introduction

The project has advanced methods for mapping, quantifying and understanding the Tanzanian savanna forest resources. The research (validated maps of forest baseline and forest change maps, as well as the carbon loss estimate) have demonstrated the benefits and limitations of optical data (high resolution RapidEye and low resolution Landsat data) for supporting REDD+ activities in the savanna forests of Tanzania. The conclusions hold the prospect of benefiting the country in its quest to move towards a sustainable use of the forest resource through promoting more sustainable utilization of these systems and nature conservation. The results also directly inform national and international policy in relation to greenhouse gas emissions and biodiversity conservation.

8.2 Lessons Learnt from Liwale Fieldwork

The fieldwork undertaken in Liwale during February 2014 was useful as field data for 133 field points was captured and these areas with known LULC were useful during the development of the rule-sets used for the LCCS classification process. In addition, the exposure to the Liwale savanna forests assisted in understanding the remote sensing images and over the area and in their interpretation.

The field observations established that much of the forests were being cleared or degraded, pri-

marily for agriculture and particularly shifting cultivation. Such conversion can contribute to soil degradation, and the progressive conversion of forests to shrublands, to grasslands (Holden, 2001). In the past, shifting cultivation was primarily for food production to sustain the local population. However, the recent cultivation of sesame plants promises better returns than other crops, namely cashew nut trees. This change of crop type leads to greater clearing compared to maize *shambas* (Figure 8.1). Furthermore, for sesame cultivation, smaller trees are completely cleared (e.g., stumps are burnt) whereas before these would be left as stumps from which regeneration would occur subsequently. A solution would be to ensure that these stumps are retained to ensure regeneration into the future.

Figure 8.1: A large forest area that had been newly cleared for sesame production.

Generally, high volumes of tree stumps resulting from forest clearing for subsistence food production

were observed. An approach that would benefit the environment is one where farmers would care for the stumps in their fields, as they hold the potential of quickly replenishing these areas that had been degraded immediately after they are abandoned. Currently, as highlighted, farmers will either prune stumps (shown with a red polygon in Figure 8.2) every growing season, or worse off, keep them under check through burning (shown with a blue polygon in Figure 8.2). Depending on the intensity of the fire and frequency of the burning, some stumps are dead by the time the *shamba* is abandoned. Community driven initiatives that would encourage farmers to practice pruning over the use of fire would have a great environmental impact in this regard. In the event a farmer feels strongly the need to use fire, the initiative could encourage them to use it minimally, and be of low intensity. The practice would, overall, improve regrowth and thus encourage forest regeneration immediately after a field abandonment. The use of the pruning approach and minimal use of fire would also ensure that there is minimal disturbance on seed banks within the cleared areas, something that would also enhance forest regrowth after the abandonment of the fields. The destruction of both stumps and the seed bank as a result of fires exposes the areas to invasion by alien vegetation. Other abandoned areas were found to be struggling to regenerate, and became grasslands.

Figure 8.2: Tree stumps clearing methods by farmers in Liwale.

The need for continued practice of shifting cultivation is partially attributed to the non-use or poor use of soil implements, which would otherwise improve and keep *shambas* productive for longer seasons than they currently are. A drive and willingness by farmers to invest on herbicides used to clear herbaceous vegetation and grasses in areas being cleared into new *shambas* was noted in Liwale (Figure 4.20). This is another crucial area that authorities and policy makers can exploit, by engaging and encouraging communities/farmers to consider shifting their financial investment from herbicides to inputs such as fertilizers and hybrid seeds. The use of inputs such as fertilizers

would improve soils and thus allowing *shambas* to remain productive for longer seasons than the current time-frame of 2-5 years. Ultimately, that would result in minimal shifting or at least lengthened seasons before a farmer felt the need to shift.

Other sub-Saharan countries such as Botswana, Zambia, Swaziland, and South Africa, even though it may have been more as a result of shortage of land, have almost entirely moved from practising shifting agriculture to intensive subsistence agriculture. In the Tanzanian setting, a change by farmers to the latter would likely yield two-fold benefits. First, it would be an environmental benefit as fields would remain productive for longer years which would result in less forest clearing at a lower frequency, and hence a reduction in degradation. Secondly, this would be a direct benefit to the farmers. Clearing of forestland every few seasons is a taxing task to farmers, requiring intensive labour. Being able to cultivate the same *shamba* that has already been cleared for longer seasons would relieve farmers of this strain.

8.3 Optical Data Availability

The reviews of both time-series dataset of RapidEye and Landsat over Liwale showed that Liwale was poorly covered by remote sensing data. The limitation is as a result of persistent cloud cover over Liwale and the region. From the review, it was noted that cloud cover occurred during both the wet and dry seasons at Liwale which limited acquisitions of useful data. This greatly limit data availability if cloud cover persists over a couple of repeats.

Liwale and savannas in general are seasonal environments and the variability in canopy cover and phenology complicate the classification of forests. This is further compromised by cloud cover as this reduces the number of observations from which to track the seasonal dynamics of vegetation. For this reason, the use of SAR or LiDAR is recommended as these data are not affected by cloud cover. However, the integration of these datasets is likely to be beneficial (Waske and Van Der Linden, 2008; Reiche et al., 2015; Lehmann et al., 2015). For example, Reiche et al. (2015) reported that both spatial and temporal accuracies for a fused Landsat-ALOS PALSAR case were consistently higher than those of the Landsat and ALOS PALSAR only cases. It holds the prospect of both improving the results and filling for missing data.

8.4 Tanzanian Savanna Forests Mapping and Monitoring

Both RapidEye and Landsat sensor data were shown to be useful for mapping and monitoring forests clearing and degradation at Liwale, with the techniques developed being transferable to other areas supporting wooded savannas. The classification scheme was initially based on the LCCS taxonomy (Anderson et al., 1976). A key criteria in the application of this method is that data are provided in units of top-of-atmosphere but ideally surface reflectance to ensure that vegetation and other spectral indices are appropriately calculated. This was achieved using the ARCSI set of libraries (Bunting et al., 2014a), for both sensors. The advantage of this approach is that the software can be run on a high performance computer, which allows correction of large amounts of image data.

Once corrected, these images can be used as input to the rule-based classification operated by the EODHAM system. Changes in land cover classes over time can also be quantified. A similar rule-based technique that used multi-temporal remotely sensed data was developed by Wondrade et al. (2014) for the Lake Hawassa Watershed in Ethiopia, which produced LULC maps with overall accuracies ranging from 82.5% to 85.0% for land cover classes. The study concluded that the achieved accuracies were reasonable, and the observed classification errors were attributable to coarse spatial resolution and pixels containing a mixture of cover types (Wondrade et al., 2014). In this study, for both the forest baselines for both the high resolution RapidEye data and the lower resolution Landsat data, overall accuracies achieved were over 90%, achieving better accuracies than Wondrade et al. (2014).

The overall accuracies in the detection of change in forest cover were also >90% for both when RapidEye and Landsat sensors. This suggested that both sensors were suitable for this purpose. However, greater detail was obtained using the higher resolution RapidEye data. Such data are needed to capture changes in an environment that is heavily populated and where there is a reliance on forest resources for needs such as energy (e.g., fuel wood and charcoal) and building (e.g., timber). Whilst the Landsat can indicate such changes, the spatial resolution is generally too coarse to detect the more subtle changes occurring. This highlights that one has to consider these factors when deciding upon the best sensor to employ.

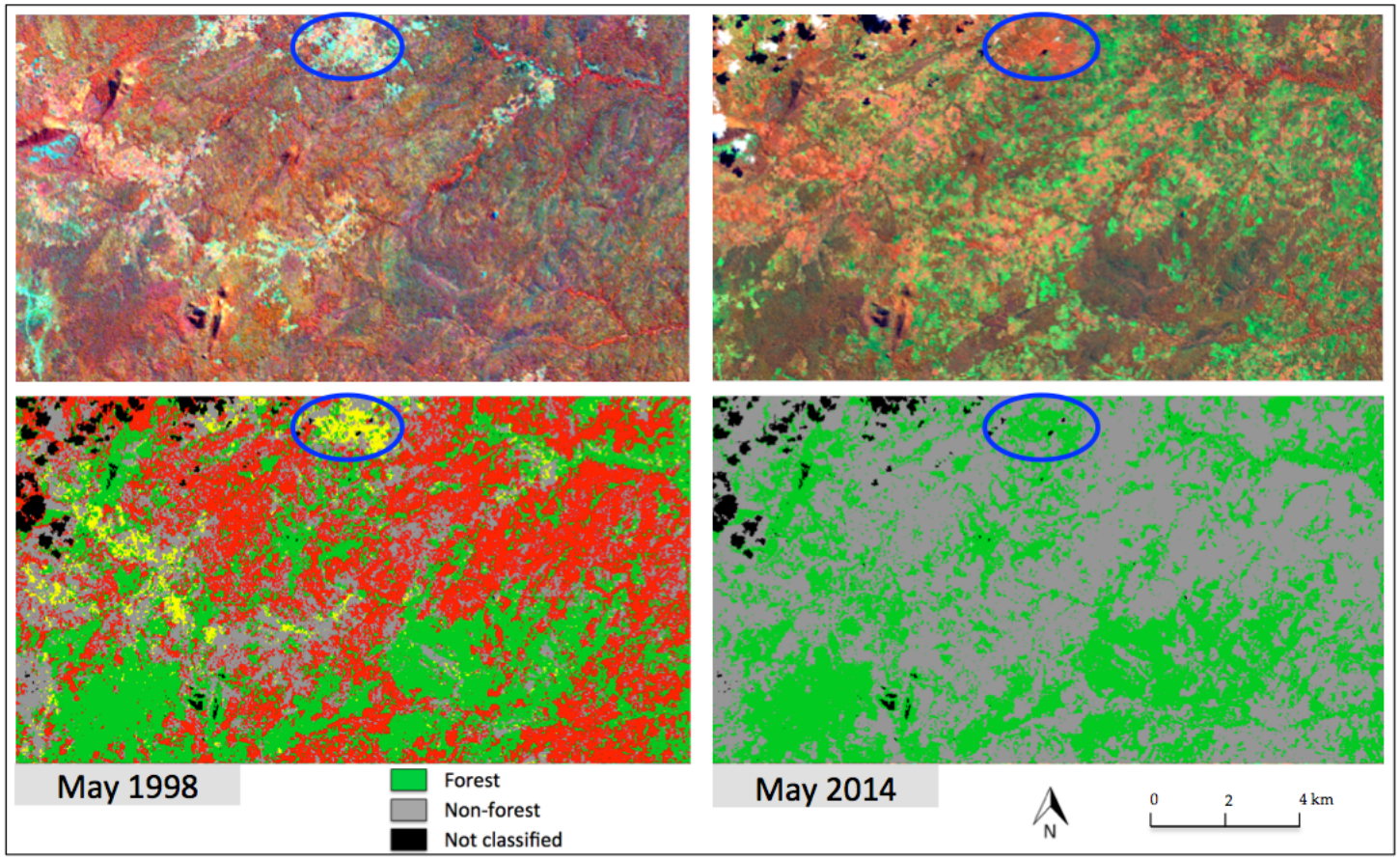


Figure 8.3: Areas of regeneration (blue) observed in time-series of Landsat sensor data (from 1998 to 2014), highlighting how shifting cultivation is shaping the forests of Liwale. The Landsat false colour images were displayed as B:NIR, G:SWIR1, R:Blue.

Using the RapidEye data from 2012 to 2014, only forest loss was detected over the short time frame. With Landsat sensor data, changes were detected over a longer period of approximately 16 years(1998 - 2014), and loss as well as gain were quantified. The change detection technique was found to be very effective for the mapping of forest gain as well. As the validation results for the change map derived from Landsat show, both users and producers accuracies of both the loss and gain classes were over 85%. A subset of Liwale shows areas of change from non-forest to forest (regrowth) between 1998 and 2014 in Figure 8.3 (yellow, circled in blue).

8.5 Forest definition for Tanzania

In the proposed monitoring approach where the forest baseline is further used in the estimation of forest change (loss and gain), limitations and the errors arising from the use of an inappropriate forest definition would further propagate into the change monitoring stage of the system. For the range of forest definitions trialled in Liwale, the results showed that the lower the forest height and the forest percentage cover used, the lower the ability of optical remote sensing data to adequately monitor the savanna forests of Tanzania. Table 5.12 shows that the adopted REDD forest definition of a REDD country determines to what extend and what ability remote sensing data can adequately and effectively map the forest baseline of the savanna environments. This was also noted by Magdon et al. (2014) that different forest definitions have different effects on the forest cover maps and estimates. The research showed that the current adopted forest definition for Tanzanian REDD+ monitoring posed limitations on the establishment of a forest baseline, using remote sensing data.

For high resolution remote sensing data (RapidEye), it was noted that the accuracy of forest/non-forest product improved greatly when the forest definition was changed from a minimum tree height of 2 m to 5 m and canopy cover of 10% to 30%. When the CHMs of the currently adopted forest definition and that of the proposed forest definition were compared with the produced forest baseline, it was also noted that there was much greater bias for former compared to the latter. Even though the overall accuracy of the comparison of the baseline to the currently adopted forest definition was slightly better than that of the forest definition being proposed (71.8% compared to 70.8%), there was too much omission of the non-forest class (commission of the forest class). However, this greatly improved when the proposed forest definition was used, even though there

was still omission, as discussed in Chapter 5. One major conclusion from the comparison of the different CHMs to the derived forest baseline was that the height attribute was more crucial than percentage cover, even though percentage cover itself does affect the results.

Figure 5.18 also highlighted the great extent the currently adopted forest definition over-classified (great commission) the forest class. Because of the very low minimum values for the forest height and canopy cover attributes, vegetated but non-forest areas were seen as forest. Examples include shrubs and and fields with crops. An increase of the minimum values of these forest attributes to those of the proposed forest baseline saw an improvement, with most of these areas and classes that had been confused being correctly classified. This improvement was in agreement with the improved level of bias as discussed above.

UN FAO highlights that the key criteria for choosing and defining a forest include that they should be unambiguous and serve the purpose, which is to assess carbon stock changes and greenhouse gas emissions and removals resulting from an activity. Furthermore, it states that parameters should be measurable during assessments (which is by remote sensing data), and should permit synergies and cost effective assessment and reporting (Schoene et al., 2007). It advocates for a forest definition that will adequately serve the forest mapping and monitoring of a country. Therefore, having shown that for the savanna woodlands of Liwale, the current adopted country forest definition for REDD monitoring limits what can be done with regard to remote sensing-based forest monitoring, it may be crucial to reconsider the definition.

For the three UNFCCC parameters for defining a forest (minimum area between 0.05-1 ha, minimum tree height of 2-5 m and minimum canopy cover of 10-30%), Tanzania adopted a forest definition for REDD monitoring that defines a forest as an area meeting the lowest criteria for all three parameters. It was found that the accuracy of separating forest from non-forest was greatly improved with an increase on these parameters. This was especially the case for the combination of tree height and canopy cover. If a national monitoring system will be based on the use of remote sensing and satellite images, there is a need to review the country forest definition therefore. With the limitation observed on high resolution data which is usually commercially available, and thus expensive, the current forest definition does not permit the establishment of a national monitoring system that utilises coarser, yet freely available data, such as that provided by Landsat. However, the proposed forest definition of minimum tree height of 5 m and minimum canopy cover of 30% was found to be appropriate and adequate for use with both high resolution and low resolution

data. The country may therefore, need to review the forest definition in order to enable remote sensing data to be used for forest monitoring, producing results of high accuracies.

The findings highlighted the importance of the use of an appropriate forest definition to ensure the adequacy and effectiveness of remote sensing data for forest monitoring in a country. The REDD programme defines the range of minimal values a country may adopt to define a forest, and further encourages the REDD countries to adjust the country definition as per need to ensure that the monitoring process is a success. The REDD+ programme is built around the use of remote sensing data for mapping and monitoring the forest resources (UNFCCC, 2002). In Tanzania where the adopted forest definition was found to be inappropriate and incapable of adequately mapping and monitoring the forest resources, it may be advisable to consider modifying the current minimal forest properties.

8.6 Limitations with LCCS Classification Approach

During the validation process, the complexity of the environment was noted, with a lot of overlapping feature space. Lucas et al. (2015) concluded that the greatest difficulty in the mapping of LCCS classes is the differentiation of aquatic vegetation as well as cultivated and managed areas, which was found to hold true in the savanna environment. Generally, the environment is highly fragmented, with the high discontinuity leading to overlaps in feature space. Furthermore, even though RapidEye data has more detail, and allows for the separation of features to a higher degree than Landsat, both sensors are limited and do not give adequate information for these environments. As shown by the overall accuracies of both RapidEye and Landsat for the environment, overall, RapidEye does not seem to give a better result than Landsat, irrespective of the difference in resolution, which was attributed to RapidEye's poor spectral properties. Even though it is a high resolution imagery, it is poor to the extent that boundaries between classes are very hard to identify.

Wetlands were spectrally similar to mature forest, and proved difficult to separate from the forest class. This led to some wetlands being classified as forest areas in the baseline. Productive forest, in many cases, was very similar spectrally to shrubs and herbs. Without the benefit of height in optical data, there were those areas where it was not possible to separate these. This was most pronounced on the RapidEye data that has enough detail for such features to be identified. For example, during

the validation process, there was productive forest areas that were incorrectly classified as non-forest. There was also areas that were non-forest (either fields with crop, herbs/closed grassland or shrubs) that were incorrectly classified as forest (productive forest). This is mainly due to the high discontinuity and heterogeneity experienced in the area. Different cover types and land use types are irregular fragments across the environment leading to a very complex and irregular feature space that becomes hard to handle.

8.7 Biomass Monitoring in Liwale

8.7.1 Introduction

Quantifying ecosystem carbon stocks is vital for understanding the relationship between changes in land use and carbon dioxide emissions (Ryan et al., 2010), and establishing the amount of carbon being lost to the atmosphere. Forests play a huge role in reducing carbon emissions as they sequester and store more carbon than any other terrestrial ecosystem (Gibbs et al., 2007). When forests are cleared or degraded, their stored carbon is released into the atmosphere as carbon dioxide. The UN FAO (2006b) reported that the largest source of greenhouse gas emissions in most tropical countries was from deforestation and forest degradation, with deforestation accounting for approximately 70% in the Africa continent.

Unfortunately, carbon cycle in African woodlands is comparatively understudied (Kalaba et al., 2013). There is relatively scarce knowledge of growth rates and wood biomass in natural woodlands, which has led to carbon stores in these woodlands, and global forests generally, uncertain (Baccini et al., 2012; Watson et al., 2000). Shifting (slash and burn) agriculture and charcoal production were reported by Chidumayo (1994) as the major causes of forest loss in Miombo woodlands, and continue to be linked to huge losses of carbon and biodiversity of forest systems (Kalaba et al., 2013).

Forest degradation is often poorly understood, yet it may be a larger source of greenhouse gas emissions than deforestation. Moreover, The Institute of Resource Assessment (2010) stated that it tend to lead to deforestation. Tanzanian forests experience disturbance in the form of both deforestation and degradation. Currently, there is lack of local scale carbon maps (estimates), except for global products (Baccini et al., 2012; Saatchi et al., 2011), and these small scale maps

have a course resolution.

8.7.2 Potential for Biomass Monitoring in the Savanna Woodlands

Tanzania is a REDD+ country that need to quantify carbon stocks and forest change on continual monitoring basis. Knowledge of biomass (carbon) estimates and the ability to monitor them is crucial for REDD+ monitoring and the successful implementation of the REDD+ framework. Studies such as Mugasha et al. (2013) and Kalaba et al. (2013) demonstrated models that are capable of estimating biomass in the savanna forests of Africa, and these studies were undertaken in Tanzania and Zambia respectively. However, the challenge with such studies lies in that, as highlighted by Houghton (2005), when a limited area is sampled and then used to predict biomass over large tracts of forest, it introduces uncertainties on the derived estimates. Moreover, it is needful to go beyond the illustration that models are applicable and scalable, to producing large-scale carbon stocks estimates.

Remote sensing technologies integrated with field measurements have a strong potential to assess national and global carbon stocks (Saatchi et al., 2011; DeFries et al., 2007; Baccini et al., 2012). However, these products, the methods used and available data also struggle to deliver the desired precision to assess and monitor forest stocks (Bustamante et al., 2015). Until better methods and data become available, these global products are the best large-scale carbon products. Having estimated the amount of forest change in Liwale, the Baccini et al. (2012) map was the most reliable and available product that could have been used to estimate the amount of carbon loss in the site. The immediate limitation with its use was that biomass associated with regrowth (forest gain) is not really quantified, and the focus was on mature forests Baccini et al. (2012).

In addition, the pixel resolution of the Baccini et al. (2012) map was 250 m, yet the change maps derived from RapidEye and Landsat were at 5 m and 30 m spatial resolution respectively. The large variance in the resolution between the biomass map and the change maps would have limited estimation of biomass loss. For the relatively sparsely forested savanna forests, the biomass map is generalised and this compromises the estimates. The different resolutions of the biomass and change maps also reduced confidence in the associated losses of carbon.

Another limitation is that the Baccini et al. (2012) carbon map was produced using ICESAT LiDAR data for the year 2003 and MODIS data for 2000s. Therefore, the acquisition time of the data used

to derive the carbon map and the acquisition time of the data used to derive the forest baseline and the forest change are different. This difference in the acquisition times would introduce errors. Furthermore, the biomass estimate loss is a value (and product) that has not been validated as there is no available field data or product that would be used for the validation process. However, the underlying conclusion is that the techniques developed for establishing a forest baseline and detecting forest change can be used to monitor the carbon stocks in the savanna forests. But there is a need for a detailed and up-to-date carbon map that would enable the estimation of forest loss/gain with less uncertainty. Field data that will be useful in estimating carbon over Liwale, and thus validate how well biomass was estimated, will be crucial.

The change maps produced from both RapidEye and Landsat show that there was forest disturbance being experienced in Liwale, with a majority occurring in and around villages. These losses were observed in an area classified as the least populated in the country, with a density of 2 people per km^2 (The United Republic of Tanzania, 2013). Between 1988 and 2012, the region of Lindi and the district of Liwale continued to experience minimum population growth (Table 3.2). In many cases, shifting cultivation was occurring in forests that were unprotected. Protected areas were found to experience none or minimal losses, with most associated with forest fires and herbivorous wildlife (e.g., elephants).

To successfully reduce emissions and meet the REDD+ requirements, an increased effort towards the adoption of more sustainable land use practices by subsistence farmers in the area, and nationally need to occur. Products such as Baccini et al. (2012) provide products that can be used together with a validated change map, to produce an up-to-date carbon baseline for monitoring, as well as the carbon credit market. The limitation with these products is their course resolution, thus the need to work towards producing a higher resolution carbon map hwich will capture more detail and better account for the heterogeneous environment of open savanna woodlands.

LiDAR data provides accurate measurement accuracy of vegetation properties such as forest height and species (Lefsky et al., 1999; Wynne, 2006; Popescu et al., 2011; Naidoo et al., 2012; Cho et al., 2012), and good estimates of forest biomass (Nelson et al., 2009; Sun et al., 2011; Clark et al., 2011; Mitchard et al., 2012; Kandel et al., 2013). The biomass loss/gain estimation in Liwale could have benefited from the use of the LiDAR data acquired over Liwale in the years 2012 and 2014. However, due to an agreement in the memorandum of understanding signed between Aberystwyth University and the Norwegian University of Life Sciences that this research could only use the

provided LiDAR data for validation purposes only, and especially not for biomass estimation, the data could not be used for this purpose. Rather, the Norwegian partners have been undertaking research aimed at producing a biomass map of the site using the LiDAR data. The success of the research will result in a local biomass map which can then be used to refine the biomass loss/gain estimation over the respective years using RapidEye and Landsat time-series data.

8.8 Current State of Tanzanian Savanna Forests from a Conservation Viewpoint

Even though amongst the least populated district with an average of two people per square kilometre (United Republic of Tanzania, 2003), Liwale has experienced substantial forest loss in the past two decades, as shown by the change maps produced from RapidEye and Landsat data. Protected areas, in the case of Liwale, the northern part of the site (Selous Game Reserve), have enjoyed minimal disturbance, and their state has remained mainly preserved. The key source of disturbance for these forests is forest fires, which may be either human-induced or natural. Savanna forests are highly flammable systems during the dry season, and are strongly influenced by anthropogenic fires (Ryan et al., 2012). Furthermore, frequent anthropogenic fires positively influence the tree species richness and create more heterogeneity within the forests (Käll, 2006). Therefore, these fires may be beneficial for protected savanna forests.

Forests adjoining or in proximity to roads and rivers are most affected. Hence, conservation efforts should focus on such areas. Whilst deforestation is widespread, the added loss of forest structure, species and biomass through degradation is less easy to quantify. Therefore, efforts need to focus on this process, particularly as substantial losses may be occurring across the region, and it is currently unaccounted for. In the quest by the country to move towards establishing sustainable forests at national level for the carbon credit market, there will be a need to not only focus on sustainable use and management of protected forest areas, but unprotected forests as well.

8.9 Proposed Components of a Monitoring System

Figure 8.4 is a simple monitoring system the research is proposing for monitoring the savanna forests of Tanzania and other similar environments in the region. The system being proposed is

made up of two major components, namely, establishing a forest baseline from a single date remote sensing data, and using a satellite image of a later date together with the established baseline to detect change (change map). From the change map, an updated forest baseline for the area of interest can then be produced. In Figure 8.4, using RapidEye satellite data, a 2012 image was used to produce a forest baseline for the year. Then a 2014 image was used to find change within the 2012 baseline, and a 2014 change map was produced. The change map was then used to produce an updated forest baseline for 2014. The system proposes therefore that the next appropriate image for 2015 or coming years is used to detect areas of change from the 2014 baseline, and thus producing the latest updated forest baseline. This becomes an ongoing process where change is detected every time an appropriate image both in terms of cloud cover and seasonality becomes available.

A common limitation with monitoring systems is data unavailability. It can be very tricky both in terms of cost and availability of appropriate data covering the whole area of interest, in this case being the country. For Tanzania, it was noted that persistent cloud cover has led to the country having very poor data coverage, as evidenced in the review of available data for both RapidEye and Landsat imagery. A system that will require that imagery covering the whole area of interest, therefore, would likely have limitations as a result of likely data shortage in this environment. The proposed system uses a very flexible classification-to-image technique for detecting change, which can be undertaken at any scale, be at individual scene level, or at a larger scale. This means that even for a single new scene (image), immediately one that is appropriate becomes available, the existing baseline can be immediately updated over that scene coverage area. Scenes can be used in singularity to update the baseline over those areas they cover, and one does not have to wait to have an imagery covering the whole area being monitored. This is very handy as different areas will always have images being available at different times (and/or years) of the wet season period. This flexibility will ensure that the baseline is updated immediately an appropriate image becomes available.

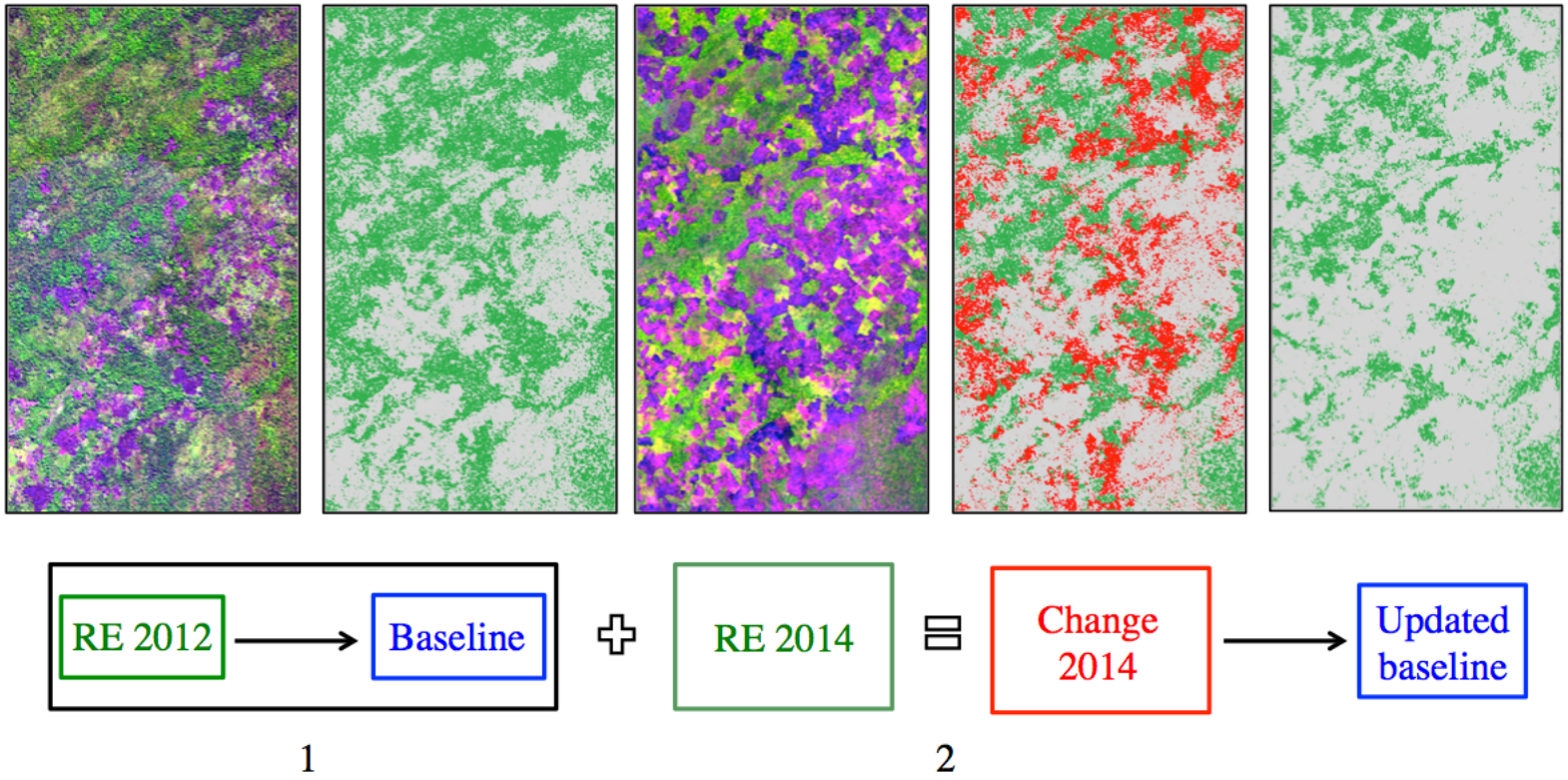


Figure 8.4: Components of the monitoring system proposed for Tanzanian savanna forests.

A monitoring system needs to be able to identify and correct for error. Hence, error accumulated during the forest classification and change detection procedures needs to be quantified. The error in the change detection process is likely to be largest and hence methods need to be developed to quantify and compensate for such errors. Figure 8.5 is a flowchart of the proposed monitoring system for these savannas. As already discussed, it is capable of producing a forest baseline, and then use newly acquired images to continually detect change on the existing baseline and produce an updated forest baseline. The new baseline is then used when a new image becomes available. Error accumulated during the production of the forest baseline, which in turn is used during the detection of change is propagated through the system. Moreover, besides errors accumulating from the baseline into the change detection stage, more error is likely to occur during change detection (shown by a red box in Figure 8.5). Hence methods need to be developed to quantify and compensate for such errors.

This is an aspect this work did not account for, with substantial research still needed to be channelled towards the development of techniques that identify errors. This is especially because literature does not account much on such techniques, understandably because there is no such a system that has been used for forest monitoring. In this work, instead, the possible sources of error were discussed. Moreover, the statistical analysis developed by Olofsson et al. (2013) and Olofsson et al. (2014) were used to assess the accuracy of products derived by the proposed system. This analysis provide accuracy measures that are accompanied by their levels of uncertainties, which is very useful in informing how well and robust the different components are. Moreover, Olofsson et al. (2013)'s stratified area estimator enabled the work to further produce area estimates of each class with a margin of error (at approximate 95% confidence interval), which was undertaken factoring in the standard error of the adjusted area estimate for each class. The good user's, producer's and overall accuracies, presented with their level of uncertainty proved the robustness of the system. Moreover, the consistency in the results for both high and lower resolution imagery shows the flexibility and the transferability of the system.

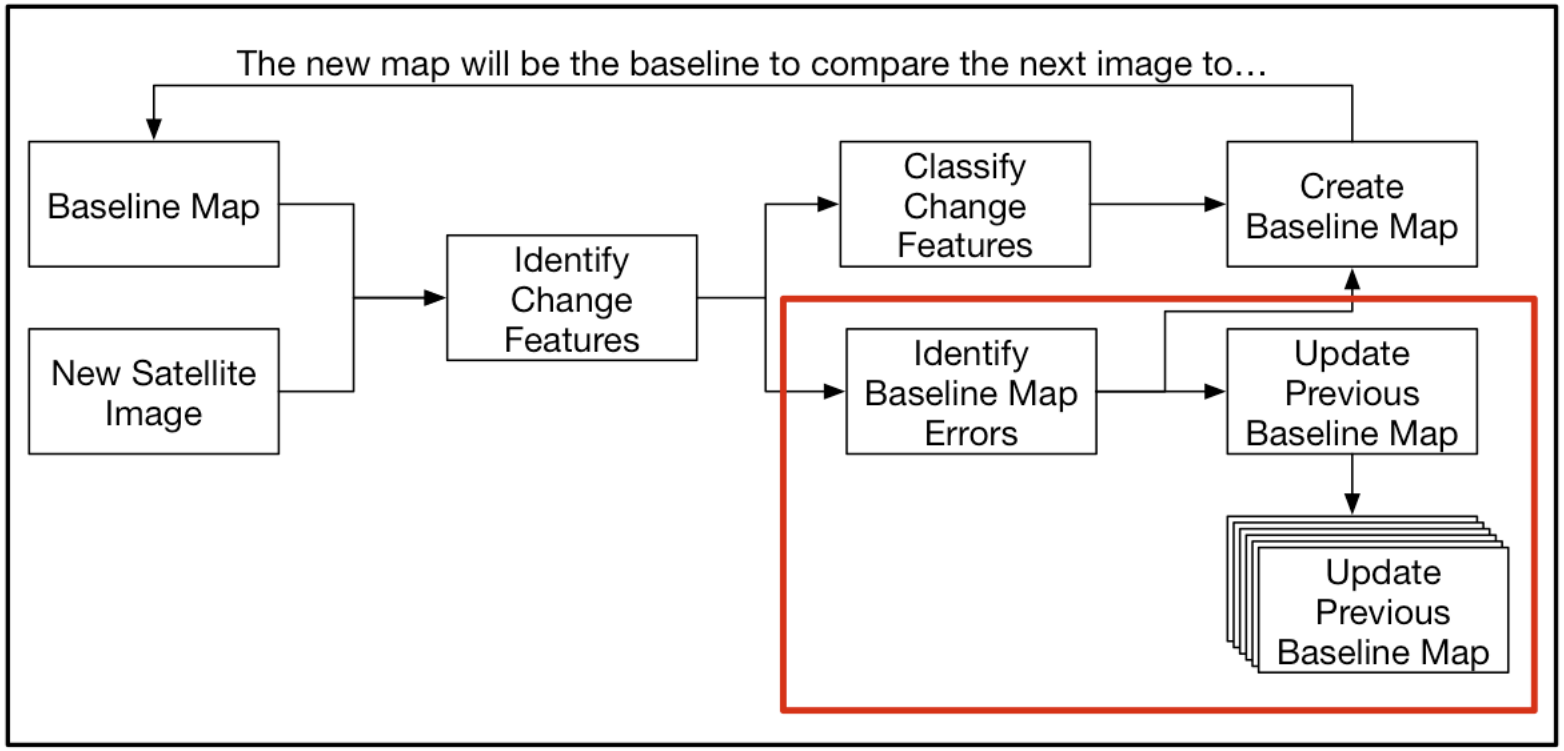


Figure 8.5: Error component (boxed in red) of the proposed monitoring system for Tanzanian savanna forests.

In terms of computational efficiency, the system was found to be very effective and efficient. The LCCS classification, being executable through the use of a python script, was noted to take a short to produce a classification. For RapidEye data, the average processing time was about 30 minutes for the 6,529 km^2 on a standard computer with 3.4 GHz Intel Core i7 processor and 16GB of RAM. For Landsat, this was higher, about 60 minutes for a single scene (34,255 km^2). However, overall, it was noted that the computation was more efficient for Landsat, probably also because segments (objects) are more detailed and smaller for the higher resolution imagery (RapidEye) compared to the Landsat that has lesser detail. This means that to run a computation for the whole country (945,087 km^2) for Landsat, for example, would require about 24 hours.

The parameterisation process of the random forests classifier, which the system uses to classify for change was also noted to be competitively efficient. It was found to not be affected by the number of trees used, as evidenced in Table 6.5. The larger the number of trees, it was noted that it did not lead to lower accuracy, and it did not lead to prolonged processing times. For both RapidEye and Landsat imagery, the site of Liwale covered by each dataset, respectively was classified for change by the classifier in less than 10 minutes.

In terms of transferability, the system was developed from the universally applicable LCCS classification, which makes it suited to be applied in any other environment. Lucas et al. (2015) showed that the classification system could be successfully applied in different environments around the globe. The proposed system is also very robust, yet very flexible, in that the same techniques were successfully applied in Liwale using two different datasets. The techniques that had been developed using the high resolution RapidEye data were successfully scaled out to an imagery with a coarser resolution (Landsat), and still produced acceptable results, both for the establishment of the forest baseline and change detection.

As literature (e.g. Alqurashi and Kumar (2013)) reported, within the machine learning techniques, random forests is amongst the easiest to implement (See Table 6.2). It was noted that only two parameters are needed in order to implement it; the number of trees and the gini index. This makes the technique favoured and efficient. Factoring that not many organisations or personnel are familiar with machine learning techniques, the most user friendly would be the most ideal, especially if it was found to output as much competent results as the other techniques, if not better. Random forests was, therefore, found to be ideal in this regard. In addition, As alluded to, it does not suffer from over-fitting, and can effectively deal with overlapping feature space, which

is common in LULC environments.

A crucial aspect that will be needed as further development of these components and the techniques for eliminating errors is pursued, will be a good time series data over the area. Another need for this process is good field data which is essential for validating the errors picked by the technique. Unfortunately, for this work, the study site of Liwale had poor data coverage and availability for both sensors (RapidEye and Landsat), and there was no field data available for the times these images used were acquired. Moreover, the development of this component will require a lot of resources in terms of time and finances. There will be a need to invest on field data (fieldwork) which will play a crucial role in further validating developed mapping and detection techniques, and especially for developing techniques for identifying errors and correcting for them through the system.

8.10 Selection of the Sensor for Monitoring in Tanzanian savannas

Liwale was found to experience both large-scale and small-scale clearing. There are areas that were completely cleared, as well as areas dominated by degradation, where only a few pixels have been disturbed. The high resolution RapidEye data was found to be effective for capturing such change, as evidenced by the high overall accuracies attained for both the forest mapping and change detection process. This makes it well suited for environments settled upon and with human activities. In this environment, subsistence farmers are likely to degrade an area they are settled in and around. However, it is a costly dataset which would require substantial resources, especially if used at a regional or national level. On the other hand, Landsat was found to be very effective too for both forest mapping and change detection. The use of both datasets may be a very effective approach, both in terms of data availability and cost. Vast stretches of forest areas without human disturbance, such as protected areas are not likely to experience small-scale change, and Landsat data would be adequate for monitoring them. A proposed approach is one that will create buffers around settled areas and for such, RapidEye data are used for monitoring. Then areas outside of these buffers without human disturbance should be monitored using the freely available Landsat data. The research has illustrated these two datasets (Landsat and RapidEye) are complementary, and for these techniques can even be used interchangeably. The proposed approach would be very

cost-effective, and would ensure that even at national level, RapidEye data are used to detect changes only in areas within a certain proximity to human settlement. The reduced cost would enhance the chances of tasking the acquisition of data over problematic areas. In addition, the flexible monitoring system would enable the monitoring and updating of any area whenever an appropriate data is received, it would enable the independent update of such areas upon receipt of new acquisitions.

8.11 Conclusions

Both RapidEye and Landsat are capable of being used as the primary sensors for forest mapping and forest monitoring in the savanna woodlands of Tanzania. As means to reduce cost at a large scale, it is recommended these two sensors be used interchangeably, depending on whether an area has human interaction or not. A proposed monitoring system for the environment is one that will first map the forest and establish its baseline. Thereafter the baseline is kept updated by continually using a single date image of a later date to find change areas, and thus update the existing baseline.

Chapter 9

Conclusions

Biomass (wood-based fuels) being the source of 90% of Tanzanian energy highlight the likely high rate of the forest resource depletion in the country. Tanzania is one of the REDD programme member countries, which are aiming at reducing carbon emissions from deforestation and degradation. A prerequisite for effectively reducing and/or eliminating forest disturbance resulting from both deforestation and degradation is knowing the forest resource and keeping check of it through its monitoring. It is crucial therefore to develop and advance methods and techniques for mapping and monitoring the resource. It entails tracking these forests and accounting for any changes within them, both the scale and dynamics. Techniques for mapping and monitoring the African savanna forests are lacking, yet it hosts a high populations that rely heavily on it for their source of livelihood. This research therefore, sought to advance methodologies for mapping and monitoring forest disturbance in Tanzanian forests which are currently lacking, through the use of remote sensing.

Literature showed that:

- Remote sensing has been recognised for the opportunities it avails for mapping and monitoring forest resources, especially the spatio-temporal aspect it avails
- Deforestation has been well research and techniques for mapping and detecting it have been well documented for the tropical rainforest where it was dominant
- Degradation has been poor researched and techniques are lacking for this kind of disturbance

- The African savanna environment has been poorly researched, and the limited work that has been done in these environment has not been effectively shared

9.1 Major Findings and Conclusions

This study sought to investigate three objectives, and below is a summary of the main findings from the investigations and analyses.

9.1.1 Objective 1

To develop a method for establishing a forest/non-forest baseline for Tanzanian open woodlands (miombo) from high resolution remote sensing data

A method based on LCCS classification for mapping the savanna forests using high resolution remote sensing data was successfully developed. This technique was found to map forest from non-forest with an overall accuracy of 95.81% and a kappa coefficient of 0.914 for the savanna forests of Liwale. When presented with 95% confidence interval, the overall accuracy was 94 $\pm$ 1.2%. A major limitation observed for the environment is limited data availability as a result of persistent cloud cover and seasonality. Other limitations included poor registration of RapidEye data, with the shift in the data introducing error, and a highly heterogeneous environment which led to overlap in the feature space of classes. However, the technique itself was found to be very effective for establishing a forest/non-forest baseline in the savanna environment. It has the capability to be transferred across to the wider savanna environment of Tanzania and the African savannas generally.

9.1.2 Objective 2

To develop a method for detecting forest change and for classifying the potential change features from high resolution remote sensing data

A method for detecting and classifying for change in the savanna forests of Liwale was developed, and it uses random forests classifier to perform the classification. Skewness and kurtosis is used to first threshold the data and thus identify change features from non-change features. Thereafter,

the potential change features are classified into either change or non-change features using the classifier. Among the major conclusions was that RapidEye data had the ability to map even small-scale forest disturbance (degradation) as the method was found to detect change with an overall accuracy of 96.77% and a kappa coefficient of 0.943. When presented with 95% confidence interval, the overall accuracy was 97%, with an uncertainty of ± 0.5%. The standard error of the adjusted area estimate was estimated at 31.90 km^2, and the final land change area estimate with a margin of error (at approximate 95% confidence interval) was 191.75 ± 63.80 km^2. This uncertainty meant that the true area of forest loss in Liwale could have been as low as 127.95 km^2 or as high as 255.55 km^2 at the 95% level of confidence. The ability to map small change features is critical in Liwale as there is a lot of subsistence farmers who are still heavily reliant on the natural environment to source a number of resources, and some of these lead to small-scale clearing.

9.1.3 Objective 3

To investigate if the developed forest baseline and change detection techniques can be used to train scaling out to coarser remote sensing imagery

The forest mapping and change detection methods that were developed using high resolution data were successfully scaled out to coarser resolution remote sensing data. These methods produced a forest baseline with an overall accuracy of 92.45% and a kappa coefficient of 0.819 and a change map of Liwale with an overall accuracy of 95.58% and a kappa coefficient of 0.900. A further statistical analysis used to validate these statistical measures showed that the baseline map had an overall accuracy of 92% , with an uncertainty of ±0.6%, while the change map had an overall accuracy of 96%, with an uncertainty of ±0.5%. The margin of error for area estimates was found to be 2611.88 ± 104.18 km^2 for the forest loss class and 940.66 ± 52.13 km^2 for the forest gain class. The results showed that Landsat data produced good results with overall accuracies comparable to those achieved from the high resolution data. However, as would be expected for a coarser image, the products derived from Landsat data did not have as much detail as those from RapidEye. Therefore, for areas experiencing small-scale forest disturbance (degradation), Landsat is not capable to capture such.

9.2 Relevance and Importance of the Work

As shown in literature, there is limited research on the African environment, and especially in the savanna woodlands. Methodologies for its mapping and monitoring are limited, if any. This work developed a simple monitoring system which has the potential to be implemented into an operational system. It is a system that has been shown to be very robust, transferable and repeatable. The first component of the system, of developing a forest baseline was developed from the FAO's LCCS classification system, makes the method to have the potential to be implemented universally, even in different environments. The second component, which focuses on detecting change using a classification-to-image approach is unique to the commonly used approaches (image-to-image or classification-to-classification), whose limitations were discussed. Even though there was limited available data, the successful development of these components, as well as the successful validation of these techniques highlight the great potential the system has for an environment usually lacking both remote sensing and field data. With the availability of good ground data, the system has the potential to further use the estimated change data to estimate biomass change (loss and gain). These components, and the conclusions are crucial as they inform the REDD monitoring programme in Tanzania, and will be useful for other African nations with savanna forests.

9.3 Further Work

To further test the techniques, it is hoped that the successfully developed techniques and methods will be scaled out to the bigger Tanzanian savanna environment and into other sub-Saharan countries (e.g. Swaziland, Mozambique, and South Africa). Moreover, in an attempt to test the ability of these techniques to be used for a national monitoring system, further work entails trialling the methods in other forest environments than savannas in Tanzania. To further test and develop the techniques, future work will also involve the trialling of them on more complex savanna woodlands.

As the techniques showed flexibility and robustness for different earth observation data types, it is believed that they can also be transferred to different environments in the country. A successful scaling out of the techniques to the different environments would avail a cost-effective opportunity for Tanzania to adopt this simple monitoring system at a national level.

As discussed in Chapter 9, SAR data was reported by numerous studies to have enhanced both forest mapping and change detection in different environments across the globe. Besides enhancing the results, the use of SAR that is not limited by the time of the day or by cloud cover, which is a limitation in Tanzania, SAR (ALOS PALSAR) data holds the prospect of being used for data gap filling in Liwale and nationally. Future work therefore aims at using it in integration with the optical data to explore if these techniques can further be improved. Furthermore, depending on availability of field data, the SAR data will be used to estimate AGB in the area, and thus establish more accurate biomass loss estimates.

As highlighted, a critical component of the proposed monitoring system for the savanna woodlands of Liwale and Tanzania that was not covered by the research is the development of methods for defining the error terms. Further work is crucial in this regard to ensure that the monitoring system is complete, capable of identifying error at any stage of the system, and especially capable of correcting for it in the new baseline as well as on the already existing forest baselines. This will entail the development of a field dataset which will be critical for training and validation purposes.

From a wider context point of view, this work hold potential to be adopted for LULC mapping. During the forest baseline establishment process, LULC classes were first classified for, before they were grouped into either forest or non-forest main classes. It is therefore worth exploring the capability of the method for general LULC mapping and monitoring purposes. Future work is aimed at trialling it out for LULC mapping and monitoring for savanna forests and other forest types in the SADC region. As Lucas et al. (2015) illustrated with case studies how the EODHaM method produced acceptable results in Italy, Wales (UK), Greece, Netherlands, India and Portugal, future work is aimed at exploring the transferability of these techniques to non-savanna environments in Africa and globally, both for forest mapping and LULC mapping.

Bibliography

Abdallah, J., Monela, G., 2007. Overview of miombo woodlands in Tanzania.

Abdel-Rahman, E. M., Van Den Berg, M., Way, M. J., Ahmed, F. B., 2009. Hand-held spectrometry for estimating thrips (Fulmekiola serrata) incidence in sugarcane.

Achard, F., Eva, H. D., Stibig, H.-j., Mayaux, P., Richards, T., Malingreau, J.-p., 2002. Determination of deforestation rates of the world's humid tropical forests. Science 297 (5583), 999–1002.

Adam, E. M., Mutanga, O., Rugege, D., Ismail, R., 2012. Discriminating the papyrus vegetation (Cyperus papyrus L.) and its co-existent species using random forest and hyperspectral data resampled to HYMAP. International Journal of Remote Sensing 33 (2), 552–569.

Adjorlolo, C., Cho, M. a., Mutanga, O., Ismail, R., 2012. Optimizing spectral resolutions for the classification of C3 and C4 grass species, using wavelengths of known absorption features. Journal of Applied Remote Sensing 6 (1).

Afify, H. A., 2011. Evaluation of change detection techniques for monitoring land-cover changes: A case study in new Burg El-Arab area. Alexandria Engineering Journal 50 (2), 187–195.

Ahmed, A. I. M. U., 2008. Underlying causes of deforestation and forest degradation in bangladesh. Tech. rep., Department of Sociology, University of Dhaka, Dhaka.

Akar, Ö., Güngör, O., 2012. Classification of multispectral images using Random Forest algorithm. Journal of Geodesy and Geoinformation 1 (2), 105–112.

Al-doski, J., Mansor, S. B., Zulhaidi, H., Shafri, M., 2013. NDVI differencing and post-classification to detect vegetation changes in Halabja City, Iraq. Journal of Applied Geology and Geophysics 1 (2), 1–10.

Al-Khudhairy, D. H. A., Caravaggi, I., Glada, S., 2005. Structural damage assessments from Ikonos data using change detection, object-oriented segmentation, and classification techniques. Photogrammetric Engineering and Remote Sensing 71 (7), 825–837.

Allen, T. R., Kupfer, J. A., 2000. Application of spherical statistics to change vector analysis of landsat data: Southern appalachian spruce - Fir forests. Remote Sensing of Environment 74 (3), 482–493.

Alphan, H., 2011. Comparing the utility of image algebra operations for characterizing landscape

changes: The case of the Mediterranean coast. Journal of Environmental Management 92 (11), 2961–2971.

Alqurashi, A. F., Kumar, L., 2013. Investigating the use of remote sensing and GIS techniques to detect land use and land cover change: A review. Advances in Remote Sensing 2 (June), 193–204.

Alvarez-Borrego, J., Martin-Atienza, B., 2013. Some statistical properties of surface slopes via remote sensing using variable reflection angle considering a non-gaussian probability density function. IEEE Geoscience and Remote Sensing Letters 10 (2), 246–250.

Anderson, J. R., Hardy, E. E., Roach, J. T., Witmer, R. E., Peck, D. L., 1976. A land use and land cover classification system for use with remote sensor data. A revision of the land use classification system as presented in U.S. Geological Survey Circular 671 964, 41.

Arav, R., Filin, S., 2012. Detection and Quantification of Morphological Changes Using Multi-Resolution Terrestrial Laser Scans. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences I-7 (September), 197–202.

Armston, J. D., aug 2009. Prediction and validation of foliage projective cover from Landsat-5 TM and Landsat-7 ETM+ imagery. Journal of Applied Remote Sensing 3 (1), 033540.

Asner, G. P., aug 2009. Automated mapping of tropical deforestation and forest degradation: CLASlite. Journal of Applied Remote Sensing 3 (1), 033543.

Asner, G. P., Knapp, D. E., Broadbent, E. N., Oliveira, P. J. C., Keller, M., Silva, J. N., 2005. Selective logging in the Brazilian Amazon. Science (New York, N.Y.) 310 (5747), 480–482.

Aune-Lundberg, L., Strand, G.-H., 2014. Comparison of variance estimation methods for use with two-dimensional systematic sampling of land use/land cover data. Environmental Modelling & Software 61, 87–97.

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., Houghton, R. A., 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Climate Change 2 (3), 182–185.

Baccini, A., Laporte, N., Goetz, S. J., Sun, M., Dong, H., oct 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. Environmental Research Letters 3 (4), 045011.

Balanda, K. P., Macgillivray, H. L., 1988. Kurtosis : A critical review. The American Statistician 42 (2), 111–119.

Balthazar, V., Vanacker, V., Lambin, E. F., aug 2012. Evaluation and parameterization of ATCOR3 topographic correction method for forest cover mapping in mountain areas. International Journal of Applied Earth Observation and Geoinformation 18, 436–450.

Ban, Y., Gong, P., Giri, C., 2015. Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. ISPRS Journal of Photogrammetry and ... 103, 1–6.

Banskota, A., Kayastha, N., Falkowski, M. J., Wulder, M. A., Froese, R. E., White, J. C., 2014. Forest monitoring using Landsat time-series data - A review. Canadian Journal of Remote Sensing 40 (5), 362–384.

Bardossy, A., Samaniego, L., 2002. Fuzzy rule-based classification of remotely sensed imagery. IEEE Transactions on Geoscience and Remote Sensing 40 (2), 362–374.

Baronti, S., Carla, R., Sigismondi, S., Alparone, L., 1994. Principal component analysis for change detection on polarimetric multitemporal SAR data.

Bater, C. W., Coops, N. C., feb 2009. Evaluating error associated with lidar-derived DEM interpolation. Computers & Geosciences 35 (2), 289–300.

Belgiu, M., Dr Gu, L., Strobl, J., 2014. Quantitative evaluation of variations in rule-based classifications of land cover in urban neighbourhoods using WorldView-2 imagery. ISPRS journal of photogrammetry and remote sensing : official publication of the International Society for Photogrammetry and Remote Sensing (ISPRS) 87 (100), 205–215.

Belgiu, M., Drgu, L., 2014. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. ISPRS Journal of Photogrammetry and Remote Sensing 96, 67–75.

Belward, A. S., Skøien, J. O., 2015. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. ISPRS Journal of Photogrammetry and Remote Sensing 103, 115–128.

Benediktsson, J., Swain, P., Ersoy, O., 1989. Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data.

Benediktsson, J. a., Sveinsson, J. R., 1997. Feature extraction for multisource data classification with artificial neural networks. International Journal of Remote Sensing 18 (4), 727–740.

Berberoglu, S., Akin, a., 2009. Assessing different remote sensing techniques to detect land use/cover changes in the eastern Mediterranean. International Journal of Applied Earth Observation and Geoinformation 11 (1), 46–53.

Biggs, R., Bohensky, E., Desanker, P., Fabricius, C., Lynam, T., Misselhorn, A., Musvoto, C., Mutale, M., Reyers, B., Scholes, R., Shikongo, S., Jaarsveld, A. V., 2004. The Southern African Millennium Ecosystem Assessment. Tech. rep., Pretoria.

BirdLife International, 2002. Important Bird Areas and potential Ramsar Sites in Africa. Cambridge.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65 (1), 2–16.

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic Object-Based Image Analysis - Towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing 87, 180–191.

Bock, M., Xofis, P., Mitchley, J., Rossner, G., Wissen, M., jul 2005. Object-oriented methods for habitat mapping at multiple scales Case studies from Northern Germany and Wye Downs, UK. Journal for Nature Conservation 13 (2-3), 75–89.

Bontemps, S., Herold, M., Kooistra, L., Van Groenestijn, A., Hartley, A., Arino, O., Moreau, I., Defourny, P., 2012. Revisiting land cover observation to address the needs of the climate modeling community. Biogeosciences 9 (6), 2145–2157.

Böttcher, H., Eisbrenner, K., Fritz, S., Kindermann, G., Kraxner, F., McCallum, I., Obersteiner, M., 2009. An assessment of monitoring requirements and costs of 'Reduced Emissions from Deforestation and Degradation'. Carbon balance and management 4, 7.

Boucher, D., Elias, P., Lininger, K., May-Tobin, C., Roquemore, S., Saxon, E., 2011. The root of the problem: what's driving tropical deforestation today? Union of Concerned Scientists, Cambridge.

Bouziani, M., Goita, K., He, D. C., 2010. Rule-based classification of a very high resolution image in an urban environment using multispectral segmentation guided by cartographic data.

Breiman, L., 1999. Using adaptive bagging to debias regressions. Tech. rep., University of California, Berkeley.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Brown, M., Gunn, S. R., Lewis, H. G., 1999. Support vector machines for optimal classification and spectral unmixing. Ecological Modelling 120 (2-3), 167–179.

Bryan, J. E., Shearman, P. L., Asner, G. P., Knapp, D. E., Aoro, G., Lokes, B., jan 2013. Extreme differences in forest degradation in Borneo: comparing practices in Sarawak, Sabah, and Brunei. PloS one 8 (7), e69679.

Bunting, P., 2015. ARCSI.

Bunting, P., Armston, J., Clewley, D., Lucas, R. M., jul 2013a. Sorted pulse data (SPD) library. Part II: A processing framework for LiDAR data from pulsed laser systems in terrestrial environments. Computers & Geosciences 56, 207–215.

Bunting, P., Armston, J., Lucas, R. M., Clewley, D., jul 2013b. Sorted pulse data (SPD) library. Part I: A generic file format for LiDAR data from pulsed laser systems in terrestrial environments. Computers & Geosciences 56, 197–206.

Bunting, P., Clewley, D., Lucas, R. M., Gillingham, S., 2014a. The Remote Sensing and GIS Software Library (RSGISLib). Computers & Geosciences 62, 216–226.

Bunting, P., Gillingham, S., aug 2013. The KEA image file format. Computers & Geosciences 57, 54–58.

Bunting, P., Lucas, R. M., Jones, J., Mabaso, S., 2014b. Pre-processing of Landsat and RapidEye Data. Tech. rep., Norwegian Space Centre, Oslo.

Burges, C. J. C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2, 121–167.

Burgess, N., Lovett, J., 2000. Plant Biodiversity Hotspots in Tanzania.

Burgess, N. D., Bahane, B., Clairs, T., Danielsen, F., Dalsgaard, S., Funder, M., Hagelberg, N., Harrison, P., Haule, C., Kabalimu, K., Kilahama, F., Kilawe, E., Lewis, S. L., Lovett, J. C., Lyatuu, G., Marshall, A. R., Meshack, C., Miles, L., a.H. Milledge, S., Munishi, P. K., Nashanda, E., Shirima, D., Swetnam, R. D., Willcock, S., Williams, A., Zahabu, E., jul 2010. Getting ready for REDD+ in Tanzania: a case study of progress and challenges. Oryx 44 (03), 339–351.

Burnett, C., Blaschke, T., 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecological Modelling 168 (3), 233–249.

Bustamante, M. M., Roitman, I., Aide, T. M., Alencar, A., Anderson, L., Aragão, L., Asner, G. P., Barlow, J., Berenguer, E., Chambers, J., Costa, M. H., Fanin, T., Ferreira, L. G., Ferreira, J. N., Keller, M., Magnusson, W. E., Morales, L., Morton, D., Ometto, J. P., Palace, M., Peres, C., Silvério, D., Trumbore, S., Vieira, I. C., 2015. Towards an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. Global Change Biology.

Campbell, B., Frost, P., Byron, N., 1996. Miombo woodlands and their use: overview and key issues. In: Campbell, B. M. (Ed.), The Miombo in Transition: Woodlands and Welfare in Africa. Center for International Forestry Research (CIFOR), Bogor, Ch. 1, pp. 1–10.

Campbell, J. B., Wynne, R. H., 2011. Introduction to Remote Sensing, 5th Edition. The Guilford Press, New York.

Camps-Valls, G., 2009. Machine learning in remote sensing data processing.

Carreiras, J. M. B., Jones, J., Lucas, R. M., Gabriel, C., 2014. Land use and land cover change dynamics across the Brazilian Amazon: Insights from extensive time-series analysis of remote sensing data. PLoS ONE 9 (8).

Carroll, M., Wooten, M., DiMiceli, C., Sohlberg, R., Kelly, M., 2016. Quantifying Surface Water Dynamics at 30 Meter Spatial Resolution in the North American High Northern Latitudes 19912011. Remote Sensing 8 (8), 622.

Chan, J. C. W., Paelinckx, D., 2008. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sensing of Environment 112 (6), 2999–3011.

Chander, G., Haque, M. O., Sampath, A., Brunn, A., Trosset, G., Hoffmann, D., Roloff, S., Thiele, M., Anderson, C., 2013. Radiometric and geometric assessment of data from the RapidEye constellation of satellites. International Journal of Remote Sensing (June), 37–41.

Chander, G., Haque, M. O., Sampath, A., Brunn, A., Trosset, G., Hoffmann, D., Roloff, S., Thiele, M., Anderson, C., Wu, W., De Pauw, E., Helldén, U., Foody, G. M., Stehman, S. V., Atkinson, P. M., Aune-Lundberg, L., Strand, G.-H., Foody, G. M., Gonçalves, L. M. S., Fonte, C., Júlio, E. N. B. S., Caetano, M., Sarmento, P., Carrão, H., Caetano, M., Stehman, S. V., Wang, J. F.,

Stein, A., Gao, B. B., Ge, Y., Zhang, J., Foody, G. M., 2009. A review of spatial sampling. International Journal of Remote Sensing 30 (20), 1–14.

Chave, J., Andalo, C., Brown, S., Cairns, M. a., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J. P., Nelson, B. W., Ogawa, H., Puig, H., Riéra, B., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 87–99.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., Perez, R., 2004. Error propagation and scaling for tropical forest biomass estimates. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 359 (February), 409–420.

Chen, G., Hay, G. J., Carvalho, L. M. T., Wulder, M. A., 2012. Object-based change detection techniques. International Journal of Remote Sensing 33 (14), 4434–4457.

Chiang, S.-s., Member, S., Chang, C.-i., Member, S., Ginsberg, I. W., 2001. Unsupervised Target Detection in Hyperspectral images using projection pursuit. IEEE Transactions on Geoscience and Remote Sensing 39 (7), 1380–1391.

Chidumayo, E., 1994. Phenology and nutrition of miombo woodland trees in Zambia. Trees 9, 67–72.

Chidumayo, E. N., 2013. Forest degradation and recovery in a miombo woodland landscape in Zambia: 22 years of observations on permanent sample plots. Forest Ecology and Management 291, 154–161.

Chiesa, F., Dere, M., Saltarelli, E., Sandbank, H., 2009. UN-REDD in Tanzania. Project on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries v1. 1. Tech. rep.

Cho, M. A., Mathieu, R., Asner, G. P., Naidoo, L., van Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R., Smit, I. P. J., Erasmus, B., 2012. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. Remote Sensing of Environment 125, 214–226.

Civco, D. L., mar 1993. Artificial neural networks for land-cover classification and mapping. International Journal of Geographical Information Systems 7 (2), 173–186.

Civco, D. L., Hurd, J. D., Wilson, E. H., Song, M., Zhang, Z., 2002. A comparison of land use and land cover change detection methods. ASPRS-ACSM Annual Conference and FIG XXII Congress, 12.

Clark, M. L., Roberts, D. a., Ewel, J. J., Clark, D. B., 2011. Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors. Remote Sensing of Environment 115 (11), 2931–2942.

Clewley, D., Bunting, P., Shepherd, J., Gillingham, S., Flood, N., Dymond, J., Lucas, R., Armston, J., Moghaddam, M., jun 2014. A Python-Based Open Source System for Geographic Object-Based Image Analysis (GEOBIA) Utilizing Raster Attribute Tables. Remote Sensing 6 (7), 6111–6135.

Cochran, W. G., 1977. Sampling Techniques.pdf.

Congalton, R. G., 1991. A Review of Assessing the Accuracy of Classification of Remotely Sensed Data A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of Environment 37, 35–46.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Lambin, E., 2004. Digital change detection methods in ecosystem monitoring: a review. International Journal of Remote Sensing 25 (9), 1565–1596.

Coppin, P. R., Bauer, M. E., 1996. Change Detection in Forest Ecosystems with Remote Sensing Digital Imagery Digital Change Detection in Temperate Forests. Remote Sensing Reviews 13, 612–625.

Cunningham, A., 2014. Satellite-based monitoring of land cover change in Wales: a multi-sensor approach. Ph.D. thesis, University of Wales, Aberystwyth.

Dai, X., Khorram, S., 1999. Remotely sensed change detection based on artificial neural networks. PE & RS- Photogrammetric Engineering and Remote ... 65 (10), 1187–1194.

Danielsen, F., Skutsch, M., Burgess, N. D., Jensen, P. M., Andrianandrasana, H., Karky, B., Lewis, R., Lovett, J. C., Massao, J., Ngaga, Y., Phartiyal, P., Poulsen, M. K., Singh, S. P., Solis, S., Sørensen, M., Tewari, A., Young, R., Zahabu, E., apr 2011. At the heart of REDD+: a role for local people in monitoring forests? Conservation Letters 4 (2), 158–167.

DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., Schlamadinger, B., de Souza, C., jun 2007. Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. Environmental Science & Policy 10 (4), 385–394.

Deilmai, B. R., Kanniah, K. D., Rasib, A. W., Ariffin, A., 2014. Comparison of pixel -based and artificial neural networks classification methods for detecting forest cover changes in Malaysia. IOP Conference Series: Earth and Environmental Science 18, 012069.

Devries, B., Verbesselt, J., Kooistra, L., Herold, M., 2015. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. Remote Sensing of Environment 161, 107–121.

Dewan, A. M., Yamaguchi, Y., 2009. Land use and land cover change in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. Applied Geography 29 (3), 390–401.

Dewees, P. A., Campbell, B. M., Katerere, Y., Sitoe, A., Cunningham, A. B., Angelsen, A., Wunder, S., 2011. Managing the miombo woodlands of Southern Africa. Tech. rep., Program on Forests, Washington.

Dhakal, A. S., Amada, T., Aniya, M., Sharma, R. R., 2002. Detection of areas associated with flood and erosion caused by a heavy rainfall using multitemporal Landsat TM data. Photogrammetric engineering and remote sensing 68 (3), 233–239.

Díaz-Uriarte, R., Alvarez de Andrés, S., 2006. Gene selection and classification of microarray data using random forest. BMC bioinformatics 7, 3.

Division of Environment, 2007. United Republic of Tanzania National Adaptation Programme of Action. Tech. Rep. January.

Division of Environment, 2012. National Report for the United Nations Conference on Sustainable Development, Rio+20. Tech. rep., Dar es Salaam.

Dixon, B., Candade, N., 2008. Multispectral landuse classification using neural networks and support vector machines: one or the other, or both? International Journal of Remote Sensing 29 (4), 1185–1206.

Dondeyne, S., Wijffels, a., Emmanuel, L. B., Deckers, J., Hermy, M., 2004. Soils and vegetation of Angai forest: Ecological insights from a participatory survey in South Eastern Tanzania. African Journal of Ecology 42, 198–207.

Dons, K., Bhattarai, S., Meilby, H., Smith-Hall, C., Panduro, T. E., 2016. Indirect approach for estimation of forest degradation in non-intact dry forest: modelling biomass loss with Tweedie distributions. Carbon Balance and Management 11 (1), 14.

Drgu, L., Csillik, O., Eisank, C., Tiede, D., 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. ISPRS Journal of Photogrammetry and Remote Sensing 88, 119–127.

Du, P., Liu, S., Gamba, P., Tan, K., Xia, J., 2012. Fusion of difference images for change detection over urban areas.

Dupuy, S., Barbe, E., Balestrat, M., feb 2012. An Object-Based Image Analysis Method for Monitoring Land Conversion by Artificial Sprawl Use of RapidEye and IRS Data. Remote Sensing 4 (12), 404–423.

Duro, D. C., Franklin, S. E., Dubé, M. G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sensing of Environment 118, 259–272.

Dusseux, P., Corpetti, T., Hubert-Moy, L., Corgne, S., 2014. Combined use of multi-temporal optical and Radar satellite images for grassland monitoring. Remote Sensing 6 (7), 6163–6182.

Edwards, K., Lukumbuzya, K., Kajembe, G., 2012. Capacity Needs Assessment of Government Institutions at Central, Regional, District and Local Levels for the Establishment and Management of a REDD+ Scheme in Tanzania. Tech. Rep. September, Ministry of Natural Resources and Tourism.

Egbert, S. L., Park, S., Price, K. P., Lee, R.-Y., Wu, J., Duane Nellis, M., dec 2002. Using conservation reserve program maps derived from satellite imagery to characterize landscape structure. Computers and Electronics in Agriculture 37 (1-3), 141–156.

Eisavi, V., Homayouni, S., Yazdi, A. M., Alimohammadi, A., 2015. Land cover mapping based

on random forest classification of multitemporal spectral and thermal images. Environmental Monitoring and Assessment 187 (5), 1–14.

Elnaggar, A. A., Noller, J. S., 2010. Application of remote sensing data and Decision Tree analysis to mapping salt-affected soils over large areas. Remote Sensing 2, 151–165.

ESA, 2012. SENTINEL - 2.

ESA, 2016. Access to Sentinel Data.
URL `https://sentinel.esa.int/web/sentinel/sentinel-data-access/access-to-sentinel-data`

ESA Earth Online, 2016. EO data distributed by ESA.
URL `https://earth.esa.int/web/guest/data-access/how-to-access-eo-data/earth-observation-da`

European Union, 2013. Commission Delegated Regulation (EU) No 1159/2013 of 12 July 2013 supplementing Regulation (EU) No 911/2010 of the European Parliament and of the Council on the European Earth monitoring programme (GMES) by establishing registration and licensing conditio. Official Journal of the European Union, L309/1–L309/6.

FAO, 2006a. Country Pasture/Forage Resource Profiles United Republic of Tanzania. Tech. rep., FAO.

FAO, 2006b. Global Forest Resources Assessment 2005. FAO Forestry Paper.

FAO, 2008. Land Cover Activities in the SADC Region. Tech. rep., GLCN, Rome.

FAO, 2010a. Global Forest Resources Assessment 2010: Country Report United Republic of Tanzania. Tech. rep., Rome.

FAO, 2010b. Global Forest Resources Assessment 2010: Key Findings. FAO Forestry Paper 106 (4), 12.

FAO, 2010c. Global Forest Resources Assessment 2010: Main Report. FAO Forestry Paper 147, 350 pp.

FAO, 2012. State of the World's Forests. FAO, Rome.

FAO, JRC, 2012. Global forest land-use change 19902005.

FAO, UNDP, UNEP, 2008. UN Collaborative Programme on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (UN-REDD). Tech. Rep. June.

FAO, UNEP, 2000. Land Cover Classification System (LCCS): Classification Concepts and User Manual. FAO Environment and Natural Resources Service Series, Rome.

FAO, UNEP, 2005. Land Cover Classification System. Classification concepts and user manual (Software version 2). FAO Environment and Natural Resources Service Series, Rome.

Fauvel, M., Benediktsson, J. A., 2008. Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profile. International Geoscience and Remote Sensing Symposium (IGARSS).

Fayad, I., Baghdadi, N., Bailly, J.-S., Barbier, N., Gond, V., Hajj, M., Fabre, F., Bourgine, B., nov 2014. Canopy Height Estimation in French Guiana with LiDAR ICESat/GLAS Data Using Principal Component Analysis and Random Forest Regressions. Remote Sensing 6 (12), 11883–11914.

Fisher, A., 2014. Cloud and cloud-shadow detection in SPOT5 HRG imagery with automated morphological feature extraction. Remote Sensing 6 (1), 776–800.

Foody, G. M., Mathur, A., 2004. Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification. Remote Sensing of Environment 93 (1-2), 107–117.

Foody, G. M., McCulloch, M. B., Yates, W. B., 1995. Classification of remotely sensed data by an artificial neural network: issues related to training data characteristics. Photogrammetric engineering and remote sensing 61 (4), 391–401.

Foody, M. G., 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment 80, 185–201.

Förster, M., Frick, A., Kleinschmit, B., 2011. Utilization of Spectral Measurements and Phenological Observations to Detect Grassland-Habitats with a RapidEye Intra-Annual Time-Series. MultiTemp, 265–267.

Friedl, M. a., Brodley, C. E., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 61 (3), 399–409.

Frost, P., 1996. The Ecology of Miombo Woodlands. Center for International Forestry Research (CIFOR), Bokor.

Fuller, D. O., 1999. Canopy phenology of some mopane and miombo woodlands in eastern Zambia. Global Ecology and Biogeography 8, 199–209.

Gallaun, H., Zanchi, G., Nabuurs, G. J., Hengeveld, G., Schardt, M., Verkerk, P. J., 2010. EU-wide maps of growing stock and above-ground biomass in forests based on remote sensing and field measurements. Forest Ecology and Management 260 (3), 252–261.

Gamanya, R., De Maeyer, P., De Dapper, M., 2009. Object-oriented change detection for the city of Harare, Zimbabwe. Expert Systems with Applications 36 (1), 571–588.

García, M., Riaño, D., Chuvieco, E., Salas, J., Danson, F. M., 2011. Multispectral and LiDAR data fusion for fuel type mapping using Support Vector Machine and decision rules. Remote Sensing of Environment 115 (6), 1369–1379.

Gebhardt, S., Huth, J., Nguyen, L. D., Kuenzer, C., 2012. A comparison of TerraSAR-X Quadpol backscattering with RapidEye multispectral vegetation indices over rice fields in the Mekong Delta , Vietnam. International Journal of Remote (June 2013), 37–41.

Gerwing, J. J., mar 2002. Degradation of forests through logging and fire in the eastern Brazilian Amazon. Forest Ecology and Management 157 (1-3), 131–141.

Gibbs, H. K., Brown, S., Niles, J. O., Foley, J. A., 2007. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. Environmental Research Letters 2 (4), 045023.

Giri, C., Ochieng, E., Tieszen, L. L., Zhu, Z., Singh, A., Loveland, T., Masek, J., Duke, N., 2011. Status and distribution of mangrove forests of the world using earth observation satellite data. Global Ecology and Biogeography 20 (1), 154–159.

Giri, C., Pengra, B., Long, J., Loveland, T. R., 2013. Next generation of global land cover characterization, mapping, and monitoring. International Journal of Applied Earth Observation and Geoinformation 25, 30–37.

GLC, 2015. GLCN: databases.

Goh, J. Y., Miettinen, J., Chia, A. S., Chew, P. T., Liew, S. C., 2013. Biomass estimation in humid tropical forest using a combination of ALOS PALSAR and Spot 5 satellite imagery. Asian Journal of Geoinformatics 13.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. International Journal of Remote Sensing 34 (7), 2607–2654.

Gopal, S., Gopal, S., Woodcock, C. E., Woodcock, C. E., 1996. Remote Sensing of Forest Change Using Artificial Neural Networks.

Gourlet-Fleury, S., Mortier, F., Fayolle, A., Baya, F., Ouédraogo, D., Bénédet, F., Picard, N., 2013. Tropical forest recovery from logging: a 24 year silvicultural experiment from Central Africa. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 368, 20120302.

Government of Tanzania, 2002. The Forest Act, 2002.

Grainger, A., 1999. Constraints on modelling the deforestation and degradation of tropical open woodlands. Global Ecology and Biogeography 8, 179–190.

Granitto, P. M., Furlanello, C., Biasioli, F., Gasperi, F., 2006. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. Chemometrics and Intelligent Laboratory Systems 83 (2), 83–90.

Groeneveld, R. A., Meeden, G., 1984. Measuring skewness and kurtosis. Journal of the Royal Statistical Society. Series D (The Statistician) 33 (4), 391–399.

Guyana Forestry Commission, 2012. Guyana Forestry Commission Guyana REDD+ Monitoring Reporting & Verification System (MRVS). Tech. rep., The Guyana Forestry Commission and Indufor.

Hadjimitsis, D. G., Themistocleous, K., Nisantzi, A., Matsas, A., 2010. The study of atmospheric

correction of satellite remotely sensed images intended for air pollution using sun-photometers (AERONET) and lidar system in Lemesos, Cyprus.

Ham, J., Chen, Y., Crawford, M., Ghosh, J., 2005. Investigation of the Random Forest Framework for Classication of Hyperspectral Data. IEEE Transactions on Geoscience and Remote Sensing 43 (3), 492–501.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Sohlberg, R., 2000. Global land cover classi cation at 1 km spatial resolution using a classi cation tree approach. International Journal of Remote Sensing 21 (6), 1331–1364.

Hansen, M. C., Loveland, T. R., 2012. A review of large area monitoring of land cover change using Landsat data. Remote Sensing of Environment 122, 66–74.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. a., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., Townshend, J. R. G., nov 2013. High-resolution global maps of 21st-century forest cover change. Science (New York, N.Y.) 342 (6160), 850–3.

Harris, N. L., Brown, S., Hagen, S. C., Saatchi, S. S., Petrova, S., Salas, W., Hansen, M. C., Potapov, P. V., Lotsch, A., jun 2012. Baseline map of carbon emissions from deforestation in tropical regions. Science (New York, N.Y.) 336 (6088), 1573–6.

Hartter, J., Lucas, C., Gaughan, A. E., Lizama Aranda, L., 2008. Detecting tropical dry forest succession in a shifting cultivation mosaic of the Yucatan Peninsula, Mexico. Applied Geography 28 (2), 134–149.

Hayes, D. J., Sader, S. A., 2001. Comparison of change detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series. Photogrammetric Engineering & Remote Sensing 67 (9), 1067–1075.

Hayes, M. M., Miller, S. N., Murphy, M. a., 2014. High-resolution landcover classification using Random Forest. Remote Sensing Letters 5 (2), 112–121.

He, C., Zhang, Q., Li, Y., Li, X., Shi, P., dec 2005. Zoning grassland protection area using remote sensing and cellular automata modelingA case study in Xilingol steppe grassland in northern China. Journal of Arid Environments 63 (4), 814–826.

Healey, S. P., Cohen, W. B., Zhiqiang, Y., Krankina, O. N., 2005. Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection. Remote Sensing of Environment 97 (3), 301–310.

Herold, M., Román-Cuesta, R., 2011. A review of methods to measure and monitor historical carbon emissions from forest degradation. Unasylva 238 62, 16–24.

Herold, M., Román-Cuesta, R. M., Mollicone, D., Hirata, Y., Van Laake, P., Asner, G. P., Souza, C., Skutsch, M., Avitabile, V., Macdicken, K., jan 2011. Options for monitoring and estimating historical carbon emissions from forest degradation in the context of REDD+. Carbon balance and management 6 (1), 13.

Hirschmugl, M., Steinegger, M., Gallaun, H., Schardt, M., 2014. Mapping Forest Degradation due to Selective Logging by Means of Time Series Analysis: Case Studies in Central Africa. Remote Sensing 6 (1), 756–775.

Hofmann, P., Strobl, J., Blaschke, T., Kux, H., 2008. Detecting informal settlements from QuickBird data in Rio de Janeiro using an object based approach. In: Blaschke, T., Lang, S., Hay, G. (Eds.), Object-Based Image Analysis SE - 29. Lecture Notes in Geoinformation and Cartography. Springer Berlin Heidelberg, pp. 531–553.

Holden, S., 2001. A century of technological change and deforestation in the Miombo woodlands of northern Zambia. In: Angelsen, A., Kaimowitz, D. (Eds.), Agricultural technologies and tropical deforestation. Ch. 14, pp. 251–269.

Horning, N., 2010. Random Forests: An algorithm for image classification and generation of continuous fields data sets. In: International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences. Hanoi.

Hosonuma, N., Herold, M., De Sy, V., De Fries, R. S., Brockhaus, M., Verchot, L., Angelsen, A., Romijn, E., dec 2012. An assessment of deforestation and forest degradation drivers in developing countries. Environmental Research Letters 7 (4), 044009.

Houghton, R., 2005. Aboveground Forest Biomass and the Global Carbon Balance. Global Change Biology 11 (6), 945–958.

Houghton, R. A., 2010. How well do we know the flux of CO2 from land-use change? Tellus, Series B: Chemical and Physical Meteorology 62 (5), 337–351.

Howarth, P. J., Wickware, G. M., 1981. Procedures for change detection using Landsat digital data. International Journal of Remote Sensing 2 (3), 277–291.

Howell, D. C., 2009. Chi-square test - Analysis of contingency tables. Women 35 (3), 28–83.

HPC Wales, 2015. HPC Wales.

Hu, H., 2010. Urban land-cover mapping with high-resolution spaceborne SAR data. Ph.D. thesis, Royal Institute of Technology.

Huang, C., Song, K., Kim, S., Townshend, J. R. G., Davis, P., Masek, J. G., Goward, S. N., 2008. Use of a dark object concept and support vector machines to automate forest cover change analysis. Remote Sensing of Environment 112 (3), 970–985.

Huang, K.-Y., 2002. The use of a newly developed algorithm of divisive hierarchical clustering for remote sensing image analysis. International Journal of Remote Sensing 23 (16), 3149–3168.

Huang, L., Fang, Y., Zuo, X., Yu, X., 2015. Automatic Change Detection Method of Multitemporal Remote Sensing Images Based on 2D-Otsu Algorithm Improved by Firefly Algorithm. Journal of Sensors, 8.

Huang, X., Jensen, J. R., 1997. A machine-learning approach to automated knowledge-base building for remote sensing image analysis with GIS data. Photogrammetric engineering and remote sensing 63 (10), 1185–1193.

Hughes, G., 1968. On the mean accuracy of statistical pattern recognizers.

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS Journal of Photogrammetry and Remote Sensing 80, 91–106.

Institute of Resource Assessment, 2009. Preparing for the REDD initiative in Tanzania: A synthesised consultative report. Tech. rep., University of Dar es Salaam.

IPCC, 2003. Intergovernmental Panel on Climate Change Good Practice Guidance for Land Use, Land-Use Change and Forestry. Institute for Global Environmental Strategies, Kanagawa.

Iqbal, M., 1993. International trade in non-wood forest products: An overview - International trade in non-wood fores.

Irons, J. R., Dwyer, J. L., Barsi, J. A., 2012. The next Landsat satellite: The Landsat Data Continuity Mission. Remote Sensing of Environment 122, 11–21.

Ismail, R., Mutanga, O., 2011. Infestation Using Classification Tree Ensembles and Shortwave Infrared Bands. International Journal of Remote Sensing 32 (15), 4249–4266.

Ismail, R., Mutanga, O., Kumar, L., 2010. Modeling the Potential Distribution of Pine Forests Susceptible to Sirex Noctilio Infestations in Mpumalanga, South Africa. Transactions in GIS 14 (5), 709–726.

Jensen, J., Qiu, F., Ji, M., 1999. Predictive modelling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. International Journal of Remote Sensing 20 (14), 2805–2822.

Jensen, J. R., 2000. Introductory digital image processing, 1st Edition. Prentice Hall, Upper Saddle River.

Jensen, J. R., 2005. Introductory Digital Image Processing: A Remote Sensing Perspective, 3rd Edition. Prentice Hall, Upper Saddle River.

Jensen, J. R., Cowen, D. J., Althausen, J. D., Narumalani, S., Weatherbee, O., 1993. An evaluation of the coastwatch change detection protocol in South Carolina. Photogrammetric Engineering and Remote Sensing 59 (6), 1039 – 1046.

Jimenez, L. O., Landgrebe, D. a., 1998. Supervised classification in high-dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews 28 (1), 39–54.

Jin, S., Homer, C., Yang, L., Xian, G., Fry, J., Danielson, P., Townsend, P. a., 2013. Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA. International Journal of Remote Sensing 34 (5), 1540–1560.

Jin, S., Sader, S. a., 2005. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. Remote Sensing of Environment 94 (3), 364–372.

Johnson, R. D., Kasischke, E. S., 1998. Change vector analysis: A technique for the multispectral monitoring of land cover and condition. International Journal of Remote Sensing 19 (3), 411–426.

Joseph, S., Murthy, M. S. R., Thomas, a. P., dec 2010. The progress on remote sensing technology in identifying tropical forest degradation: a synthesis of the present knowledge and future perspectives. Environmental Earth Sciences 64 (3), 731–741.

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T. A., Reiche, J., Ryan, C. M., Waske, B., 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. Remote Sensing 8 (1), 1–23.

Joyce, K. E., Belliss, S. E., Samsonov, S. V., McNeill, S. J., Glassey, P. J., 2009. A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. Progress in Physical Geography 33 (2), 183–207.

Jubanski, J., Ballhorn, U., Kronseder, K., Franke, J., Siegert, F., 2013. Detection of large aboveground biomass variability in lowland forest ecosystems by airborne LiDAR. Biogeosciences 10 (6), 3917–3930.

Kahya, O., Bayram, B., Reis, S., 2010. Land cover classification with an expert system approach using Landsat ETM imagery: A case study of Trabzon. Environmental Monitoring and Assessment 160 (1-4), 431–438.

Kaijage, E., Kuhanwa, Z., 2013. Tanzania Mapping REDD+ Finance Flows 2009-2012. Tech. rep., Forest Trends REDDX.

Kalaba, F. K., Quinn, C. H., Dougill, A. J., Vinya, R., 2013. Floristic composition, species diversity and carbon storage in charcoal and agriculture fallows and management implications in Miombo woodlands of Zambia. Forest Ecology and Management 304, 99–109.

Käll, K., 2006. The role of fire in the Miombo forest: And the adaptation of the Community-based forest management. Divaportalorg.

Kandare, K., 2011. The time series change detection methods of remote sensing. Isprs-Studentconsortium, 1–3.

Kandel, P. N., Awasthi, K., Shrestha, S. M., Hawkes, M., Kauranne, T., Gautam, B., Gunia, K., Dinerstein, E., 2013. Monitoring Aboveground Forest Biomass: A Comparison of Cost and Accuracy between LiDAR-Assisted Multisource Programme (LAMP) and Field-based Forest Resource Assessment (FRA) in Nepal. In: International Conference on Forest, People and Climate: Changing Paradigm. No. August.

Kanke, Y., Raun, W., Solie, J., Stone, M., Taylor, R., 2012. Red Edge As a Potential Index for Detecting Differences in Plant Nitrogen Status in Winter Wheat. Journal of Plant Nutrition 35 (10), 1526–1541.

Karlsson, K. G., Johansson, E., Devasthale, A., 2015. Advancing the uncertainty characterisation of cloud masking in passive satellite imagery: Probabilistic formulations for NOAA AVHRR data. Remote Sensing of Environment 158, 126–139.

Kartikeyan, B., Gopalakrishna, B., Kalubarme, M. H., Majumder, K. L., 1994. Contextual techniques for classification of high and low resolution remote sensing data. International Journal of Remote Sensing 15, 1037–1051.

Kavzoglu, T., Mather, P. M., 2003. The use of backpropagating artificial neural networks in land cover classification. International Journal of Remote Sensing 24 (23).

Kepner, W., Watts, C., 2000. A Landscape Approach for Detecting and Evaluating Change in a Semi-Arid EnvironmentInvasive Plants and FIre in the Deserts of North America. Environmental Monitoring and Assessment 64, 179–195 ST – A Landscape Approach for Detecting a.

Kim, S., McGaughey, R. J., Andersen, H.-E., Schreuder, G., aug 2009. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. Remote Sensing of Environment 113 (8), 1575–1586.

Kissinger, G., Herold, M., De Sy, V., 2012. Drivers of deforestation and forest degradation. A synthesis report for REDD+ Policymakers.

Kitabu, G., dec 2014. Probe into forest mafia syndicate down south.

Köhlin, G., Sills, E. O., Pattanayak, S. K., Wilfong, C., 2012. Energy, Gender and Development What are the Linkages? Where is the Evidence? Energy, Gender and Development.

Kotsiantis, S. B., 2007. Supervised Machine Learning: A Review of Classification Techniques. Informática 31, 249–268.

Krystyna, J., Da, N., Tokarczyk, P. A., 2009. Analysis of the Geometric Quality of the LPIS-based RapidEye level 3A product.

Kuemmerle, T., Radeloff, V. C., Perzanowski, K., Hostert, P., 2006. Cross-border comparison of land cover and landscape pattern in Eastern Europe using a hybrid classification technique. Remote Sensing of Environment 103 (4), 449–464.

Kulkarni, A. D., Lowe, B., 2016. Random Forest for Land Cover Classification. International Journal on Recent and Innovation Trends in Computing and Communication 4 (3), 58–63.

Kushwaha, S., Mukhopadhyay, S., 2010. Sustainable development planning in Pathri Rao subwatershed using geospatial techniques. Current Science (00113891), 1479–1486.

Kweka, D., Carmenta, R., Hyle, M., Mustalahti, I., Dokken, T., 2015. The context of REDD+ in Tanzania: Drivers, agents and institutions.

Lambin, E., 1999. Monitoring Forest Degradation in Tropical Regions by Remote Sensing: Some Methodological Issues. Global ecology and biogeography 8 (3), 191–198.

Lambin, E. F., Turner, B., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fischer, G., Folke, C., George, P., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E. F., Mortimore, M., Ramakrishnan, P., Richards, J. F., Skånes, H., Steffen, W., Stone, G. D., Svedin, U., Veldkamp, T. A., Vogel, C., Xu, J., dec 2001. The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change 11 (4), 261–269.

Lardeux, C., Frison, P. L., Tison, C., Souyris, J. C., Stoll, B., Fruneau, B., Rudant, J. P., 2009. Support vector machine for multifrequency SAR polarimetric data classification.

Lawrence, R. L., Wood, S. D., Sheley, R. L., 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). Remote Sensing of Environment 100 (3), 356–362.

Lawton, R. M., 1980. Browse in miombo woodland.

Lee, S.-h., Kim, C.-m., Rho, D.-k., 2001. Ecological Change Detection of Burnt Forest Area using Multi-temporal Landsat TM Data. In: ESRI User Conference. ESRI, pp. 1–8.

Lefsky, M., Cohen, W., Acker, S., Parker, G., Spies, T., Harding, D., dec 1999. Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. Remote Sensing of Environment 70 (3), 339–361.

Lehmann, E. A., Caccetta, P., Lowell, K., Mitchell, A., Zhou, Z.-S., Held, A., Milne, T., Tapley, I., jan 2015. SAR and optical remote sensing: Assessment of complementarity and interoperability in the context of a large-scale operational forest monitoring system. Remote Sensing of Environment 156, 335–348.

Lei, L., Koide, M., Iikura, Y., Yokohama, R., 1998. Correction of atmospheric effects on AVHRR imagery by 6S code. Journal of The Remote Sensing Society of Japan 18 (2), 149–162.

Lewinski, S., Bochenek, Z., 2008. Rule-based classification of SPOT imagery using object-oriented approach for detailed land cover mapping. EARSeL Symposium Remote Sensing for a Changing Europe (June), 1–8.

Lewis, S. L., Lopez-Gonzalez, G., Sonké, B., Affum-Baffoe, K., Baker, T. R., Ojo, L. O., Phillips, O. L., Reitsma, J. M., White, L., Comiskey, J. a., Djuikouo K, M.-N., Ewango, C. E. N., Feldpausch, T. R., Hamilton, A. C., Gloor, M., Hart, T., Hladik, A., Lloyd, J., Lovett, J. C., Makana, J.-R., Malhi, Y., Mbago, F. M., Ndangalasi, H. J., Peacock, J., Peh, K. S.-H., Sheil, D., Sunderland, T., Swaine, M. D., Taplin, J., Taylor, D., Thomas, S. C., Votere, R., Wöll, H., feb 2009. Increasing carbon storage in intact African tropical forests. Nature 457 (7232), 1003–6.

Li, D., Shan, J., Gong, J., 2009. Geospatial Technology for Earth Observation. Earth and Environmental Science. Springer US.

Li, X., Yeh, a. G. O., 1998. Principal component analysis of stacked multi-temporal images for the monitoring of rapid urban expansion in the Pearl River Delta. International Journal of Remote Sensing 19 (8), 1501–1518.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2 (December), 18–22.

Lillesand, T. M., Kiefer, R. W., Chipman, J. W., 2008. Remote Sensing and Image Interpretation, 7th Edition. John Wiley & Sons, Inc, Hoboken.

Lippitt, C., Rogan, J., Li, Z., 2008. Mapping selective logging in mixed deciduous forest: a compar-

ison of machine learning algorithms. Photogrammetric Engineering and Remote Sensing 74 (10), 1201–1211.

Liu, C.-H., 2004. A Model to Correct the Atmospheric Effect forSPOT/HRV Bands. Journal of Photogrammetry and Remote Sensing 9 (3), 15–30.

Liu, J., Li, W., Tian, Y., 1991. Automatic thresholding of gray-level pictures using two-dimensional Otsu method. In: International Conference on Circuits and Systems. IEEE, Shenzhen, pp. 325–327.

Liu, Q., Wu, G., Chen, J., Zhou, G., 2012. Interpretation Artificial Neural Network in Remote Sensing Image Classification. Remote Sensing, Environment and Transportation Engineering, 1–5.

Liu, X., Lathrop, R. G., 2002. Urban change detection based on an artificial neural network. International Journal of Remote Sensing 23 (12), 2513–2518.

Loveland, T. R., Dwyer, J. L., 2012. Landsat: Building a strong future. Remote Sensing of Environment 122 (October 2000), 22–29.

Lu, D., 2006. The potential and challenge of remote sensingbased biomass estimation. International Journal of Remote Sensing 27 (7), 1297–1328.

Lu, D., Mausel, P., Batistella, M., Moran, E., 2005. Landcover binary change detection methods for use in the moist tropical region of the Amazon: a comparative study. International Journal of Remote Sensing 26 (1), 101–114.

Lu, D., Mausel, P., Brondizio, E., Moran, E., 2002. Assessment of Atmospheric Correction Methods for Landsat TM Data Applicable to Amazon Basin LBA Research. International Journal of Remote Sensing 23 (13), 2651–2671.

Lu, D., Mausel, P., Brondízio, E., Moran, E., jun 2004. Change detection techniques. International Journal of Remote Sensing 25 (12), 2365–2401.

Lucas, R., Blonda, P., Bunting, P., Jones, G., Inglada, J., Arias, M., Kosmidou, V., Petrou, Z. I., Manakos, I., Adamo, M., Charnock, R., Tarantino, C., Mücher, C. a., Jongman, R. H., Kramer, H., Arvor, D., Honrado, J. P., Mairota, P., 2015. The Earth Observation Data for Habitat Monitoring (EODHaM) system. International Journal of Applied Earth Observation and Geoinformation 37, 17–28.

Lucas, R., Medcalf, K., Brown, A., Bunting, P., Breyer, J., Clewley, D., Keyworth, S., Blackmore, P., 2011. Updating the Phase 1 habitat map of Wales, UK, using satellite sensor data. ISPRS Journal of Photogrammetry and Remote Sensing 66 (1), 81–102.

Lucas, R., Rowlands, A., Brown, A., Keyworth, S., Bunting, P., 2007. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. ISPRS Journal of Photogrammetry and Remote Sensing 62 (3), 165–185.

Lucas, R. M., Armston, J., Carreiras, J., Clewley, D., Bunting, P., 2010. Use of ALOS PALSAR data for Quantifying the Biomass and Structure of Wooded Savannas, Queensland, Australia.

Lucas, R. M., Clewley, D., Accad, A., Butler, D., Armston, J., Bowen, M., Bunting, P., Carreiras, J., Dwyer, J., Eyre, T., Kelly, A., McAlpine, C., Pollock, S., Seabrook, L., 2014. Mapping forest growth and degradation stage in the Brigalow Belt Bioregion of Australia through integration of ALOS PALSAR and Landsat-derived foliage projective cover data. Remote Sensing of Environment 155, 42–57.

Lucas, R. M., Honzak, M., Foody, G. M., Curran, P. J., Corves, C., 1993. Characterizing tropical secondary forests using multi-temporal Landsat sensor imagery. International Journal of Remote Sensing 14 (16), 3061–3067.

Luoga, E. J., Witkowski, E. T. F., Balkwill, K., 2000. Subsistence use of wood products and shifting cultivation within a miombo woodland of eastern Tanzania, with some notes on commercial uses. South African Journal of Botany 66, 72–85.

Magdon, P., Fischer, C., Fuchs, H., Kleinn, C., 2014. Translating criteria of international forest definitions into remote sensing image analysis. Remote Sensing of Environment 149, 252–262.

Magdon, P., Kleinn, C., 2013. Uncertainties of forest area estimates caused by the minimum crown cover criterion - a scale issue relevant to forest cover monitoring. Environmental monitoring and assessment 185 (6), 5345–60.

Malatesta, L., Attorre, F., Vitale, M., 2013. Vegetation mapping from high-resolution satellite images in the heterogeneous arid environments of Socotra Island (Yemen). Journal of Applied Remote Sensing 7 (1), 073527.

Malele, I. I., 2011. Fifty years of tsetse control in Tanzania: Challenges and prospects for the future. Tanzania Journal of Health Research 13 (December), 1–10.

Malele, I. I., Magwisha, H. B., Nyingilili, H. S., Mamiro, K. a., Rukambile, E. J., Daffa, J. W., Lyaruu, E. a., Kapange, L. a., Kasilagila, G. K., Lwitiko, N. K., Msami, H. M., Kimbita, E. N., 2011. Multiple Trypanosoma infections are common amongst Glossina species in the new farming areas of Rufiji district, Tanzania. Parasites & Vectors 4, 217.

Manandhar, R., Odeh, I. O. a., Ancev, T., 2009. Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement. Remote Sensing 1 (3), 330–344.

Mancino, G., Nolè, A., Ripullone, F., Ferrara, A., 2014. Landsat TM imagery and NDVI differencing to detect vegetation change: Assessing natural forest expansion in Basilicata, southern Italy. IForest 7 (2), 75–84.

Mantero, P., Moser, G., Serpico, S., 2005. Partially Supervised classification of remote sensing images through SVM-based probability density estimation.

Markham, B. L., Helder, D. L., 2012. Forty-year calibrated record of earth-reflected radiance from Landsat: A review. Remote Sensing of Environment 122, 30–40.

Markham, B. L., Markham, K. J., Thome, K. J., Barsi, J. A., Kaita, E., Helder, D. L., Barker, J. L., Scaramuzza, P. L., 2004. Landsat-7 ETM+ radiometric stability and absolute calibration. IEEE Transactions on Geoscience and Remote Sensing 42 (12), 2810–2820.

Mas, J. F., 1999. Monitoring land-cover changes: A comparison of change detection techniques. International Journal of Remote Sensing 20 (1), 139–152.

Mas, J. F., 2004. Mapping land use/cover in a tropical coastal area using satellite sensor data, GIS and artificial neural networks. Estuarine, Coastal and Shelf Science 59 (2), 219–230.

Masek, J. G., Vermote, E. F., Saleous, N., Wolfe, R., Hall, F. G., Huemmrich, F., Gao, F., Kutler, J., Lim, T. K., 2013. LEDAPS Landsat Calibration, Reflectance, Atmospheric Correction Preprocessing Code, Version 2.

Matricardi, E. A., Skole, D. L., Pedlowski, M. A., Chomentowski, W., Fernandes, L. C., may 2010. Assessment of tropical forest degradation by selective logging and fire using Landsat imagery. Remote Sensing of Environment 114 (5), 1117–1129.

Maxwell, S., 2004. Filling Landsat ETM+ SLC-off Gaps Using a Segmentation Model Approach. Photogrammetric Engineering & Remote Sensing (October), 1109–1111.

McDermid, G. J., Linke, J., Pape, A. D., Laskin, D. N., McLane, A. J., Franklin, S. E., oct 2008. Object-based approaches to change analysis and thematic map update: challenges and limitations. Canadian Journal of Remote Sensing 34 (5), 462–466.

Midgley, J., Lawes, M., Chamaillé-Jammes, S., 2010. Savanna woody plant dynamics: the role of fire and herbivory, separately and synergistically. Australian Journal of Botany 58 (19), 1–11.

Miettinen, J., Shi, C., Liew, S. C., 2011. Influence of peatland and land cover distribution on fire regimes in insular Southeast Asia. Regional Environmental Change 11 (1), 191–201.

Miettinen, J., Stibig, H.-J., Achard, F., aug 2014. Remote sensing of forest degradation in Southeast Asia-Aiming for a regional view through 530m satellite data. Global Ecology and Conservation 2, 24–36.

Miles, L., Kabalimu, K., Bahane, B., 2009. Carbon, Biodiversity and Ecosystem Services: Exploring Co-benefits: Tanzania. Tech. rep., UNEP World Conservation Monitoring Centre.

Min, S.-H., Lee, J., Han, I., 2006. Hybrid genetic algorithms and support vector machines for bankruptcy prediction. Expert Systems with Applications 31 (3), 652–660.

Mishra, N. B., Crews, K. a., feb 2014. Mapping vegetation morphology types in a dry savanna ecosystem: integrating hierarchical object-based image analysis with Random Forest. International Journal of Remote Sensing 35 (3), 1175–1198.

Mitawa, G. M., Marandu, W. Y. F., 1995. Tanzania: Country Report to the FAO International Technical Conference on Plant Genetic Resources. Tech. Rep. December 1995, FAO, Dar es Salaam.

Mitchard, E., Saatchi, S., Lewis, S., Feldpausch, T., Woodhouse, I., Sonké, B., Rowland, C., Meir, P., nov 2011. Measuring biomass changes due to woody encroachment and deforestation/degradation in a forestsavanna boundary region of central Africa using multi-temporal L-band radar backscatter. Remote Sensing of Environment 115 (11), 2861–2873.

Mitchard, E. T. A., Flintrop, C. M., jan 2013. Woody encroachment and forest degradation in sub-Saharan Africa's woodlands and savannas 1982-2006. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 368 (1625), 20120406.

Mitchard, E. T. A., Saatchi, S. S., White, L. J. T., Abernethy, K. A., Jeffery, K. J., Lewis, S. L., Collins, M., Lefsky, M. A., Leal, M. E., Woodhouse, I. H., Meir, P., jan 2012. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: overcoming problems of high biomass and persistent cloud. Biogeosciences 9 (1), 179–191.

Mniwasa, B. E., Shauri, V., 2001. Review of the Decentralization Process and it's Impact on Environmental and Natural Resources Management in Tanzania. Lawyers' Environmental Action Team.

Moghaddam, M., Pierce, L., Lucas, R., nov 2005. Radar backscattering model for multilayer mixed-species forests.

Mohammadi, J., Shataee, S., jul 2010. Possibility investigation of tree diversity mapping using Landsat ETM+ data in the Hyrcanian forests of Iran. Remote Sensing of Environment 114 (7), 1504–1512.

Mollicone, D., Achard, F., Federici, S., Eva, H. D., Grassi, G., Belward, A., Raes, F., Seufert, G., Stibig, H.-J., Matteucci, G., Schulze, E.-D., 2007. An incentive mechanism for reducing emissions from conversion of intact and non-intact forests. Climatic Change 83 (4), 477–493.

Moran, E. F., Brondizio, E., Mausel, P., Wu, Y., 1994. Integrating Amazonian Vegetation, Land-Use, and Satellite Data. BioScience 44 (5), 329–338.

Moughal, T. A., 2013. Hyperspectral image classification using Support Vector Machine. Journal of Physics: Conference Series 439, 012042.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66 (3), 247–259.

Mpingo Conservation and Development Initiative, 2013. Miombo Woodland.

Mugasha, W. A., Eid, T., Bollandsås, O. M., Malimbwi, R. E., Chamshama, S. A. O., Zahabu, E., Katani, J. Z., dec 2013. Allometric models for prediction of above- and belowground biomass of trees in the miombo woodlands of Tanzania. Forest Ecology and Management 310, 87–101.

Mukama, K., Mustalahti, I., Zahabu, E., 2012. Participatory Forest Carbon Assessment and REDD+: Learning from Tanzania. International Journal of Forestry Research 2012, 1–14.

Mukama, K. M., 2010. Participatory forest carbon assessment in Angai village land forest reserve in liwale district, Lindi region, Tanzania. Ph.D. thesis, Sokoine University of Agriculture.

Munishi, P. K. T., Shear, T. H., 2004. Carbon storage in afromontane rain forests of the eastern arc mountains of Tanzania. Journal of Tropical Forest Science 16 (1), 78–93.

Munyati, C., 2000. Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. International Journal of Remote Sensing 21 (9), 1787–1806.

Mustalahti, I., Bolin, A., Boyd, E., Paavola, J., 2012. Can REDD+ reconcile local priorities and needs with global mitigation benefits? Lessons from Angai Forest, Tanzania. Ecology and Society 17 (1).

Mutanga, O., Adam, E., Cho, M. A., 2012. High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm. International Journal of Applied Earth Observation and Geoinformation 18 (1), 399–406.

Mwampamba, T. H., 2009. Forest Recovery and Carbon Sequestration Under Shifting Cultivation in the Eastern Arc Mountains, Tanzania: Landscape and Landuse Effects. Ph.D. thesis, University of California Davis.

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment 115 (5), 1145–1161.

Næsset, E., jan 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. Remote Sensing of Environment 113 (1), 148–159.

Næsset, E., Økland, T., 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. Remote Sensing of Environment 79, 105–115.

Naidoo, L., Cho, M. a., Mathieu, R., Asner, G., 2012. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. ISPRS Journal of Photogrammetry and Remote Sensing 69, 167–179.

Nandy, S., Kushwaha, S., Dadhwal, V., mar 2011. Forest degradation assessment in the upper catchment of the river Tons using remote sensing and GIS. Ecological Indicators 11 (2), 509–513.

NASA, 2014. LandSAT Science.

Nelson, R., Ranson, K., Sun, G., Kimes, D., Kharuk, V., Montesano, P., mar 2009. Estimating Siberian timber volume using MODIS and ICESat/GLAS. Remote Sensing of Environment 113 (3), 691–701.

Nemmour, H., Chibani, Y., 2006. Multiple support vector machines for land cover change detection: An application for mapping urban extensions. ISPRS Journal of Photogrammetry and Remote Sensing 61 (2), 125–133.

Niemeyer, I., Nussbaum, S., 2006. Change Detection: The Potential for Nuclear Safeguards. In: Avenhaus, R., Kyriakopoulos, N., Richard, M., Stein, G. (Eds.), Verifying Treaty Compliance SE - 15. Springer Berlin Heidelberg, pp. 335–348.

Niemeyer, I., Nussbaum, S., 2007. Change detection using the object features.

Nitze, I., Schulthess, U., Asche, H., 2012. Comparison of machine learning algorithms random

forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. Proceedings of the 4th GEOBIA, 35–40.

Niu, X., Ban, Y., 2010. Multitemporal RADARSAT-2 Polarimetric SAR Data for Urban Land Cover Classification Using Support Vector Machine. Remote Sensing for Science, Education, and Natural and Cultural Heritage, 581–8.

Norjamäki, I., Tokola, T., 2007. Comparison of atmospheric correction methods in mapping timber volume with multitemporal Landsat images in Kainuu, Finland. Photogrammetric Engineering and Remote Sensing (February), 155–163.

Oliveira, P. J. C., Asner, G. P., Knapp, D. E., Almeyda, A., Galvan-Gildemeister, R., Keene, S., Raybin, R. F., Smith, R. C., 2007. Land-Use allocation protects the Peruvian Amazon. Science 317 (5842), 1233–1236.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., Wulder, M. A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment 148, 42–57.

Olofsson, P., Foody, G. M., Stehman, S. V., Woodcock, C. E., 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment 129, 122–131.

Olsen, K. H., Fenhann, J., 2009. A reformed CDM - Including new mechanisms for sustainable development.

Ørka, H. O., Næsset, E., Bollandsås, O. M., jun 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. Remote Sensing of Environment 113 (6), 1163–1174.

Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man and Cybernetics 9 (1), 62 – 66.

Otukei, J. R., Blaschke, T., 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. International Journal of Applied Earth Observation and Geoinformation 12 (SUPPL. 1), 27–31.

Paavola, J., nov 2008. Livelihoods, vulnerability and adaptation to climate change in Morogoro, Tanzania. Environmental Science & Policy 11 (7), 642–654.

Pal, M., 2005. Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26 (1), 217–222.

Pal, M., Mather, P. M., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment 86 (4), 554–565.

Pan, Y., Birdsey, R. a., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. a., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, a. D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., aug 2011. A large and persistent carbon sink in the world's forests. Science (New York, N.Y.) 333 (6045), 988–93.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn : Machine Learning in Python. Journal of Machine Learning Research 12 (Oct), 2825–2830.

Percy, K. E. P., Jandl, R. J., Hall, J. P. H., Lavigne, M. L., 2003. The role of forests in carbon cycles, sequestration, and storage: Forests and the global carbon cycle (Sources and sinks). Natural Resources Canada (1), 2–6.

Petit, C. C., Lambin, E. F., 2001. Integration of multi-source remote sensing data for land cover change detection. International Journal of Geographical Information Science 15 (8), 785–803.

Platt, R. V., Rapoza, L., 2008. An evaluation of an object-oriented paradigm for land use/land cover classification. The Professional Geographer 60 (1), 80–100.

Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J. C., Trianni, G., 2009. Recent advances in techniques for hyperspectral image processing. Remote Sensing of Environment 113 (SUPPL. 1), S110–S122.

Podeh, S. S., Oladi, J., Pormajidian, M. R., Zadeh, M. M., 2009. Forest change detection in the north of iran using TM/ETM+ Imagery. Asian Journal of Applied Sciences 2 (6), 464–474.

Pons, X., Pesquer, L., Cristóbal, J., González-Guerrero, O., 2014. Automatic and improved radiometric correction of landsat imageryusing reference values from MODIS surface reflectance images. International Journal of Applied Earth Observation and Geoinformation 33 (1), 243–254.

Popescu, S. C., Zhao, K., Neuenschwander, A., Lin, C., 2011. Satellite lidar vs. small footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. Remote Sensing of Environment 115 (11), 2786–2797.

Potapov, P., Yaroshenko, A., Turubanova, S., Dubinin, M., Laestadius, L., Thies, C., Aksenov, D., Egorov, A., Yesipova, Y., Glushkov, I., Karpachevskiy, M., Kostikova, A., Manisha, A., Tsybikova, E., 2008. Mapping the World's Intact Forest Landscapes by Remote Sensing. Ecology and Society 13 (2), 51.

Pradhan, R., Ghose, M. K., Jeyaram, A., 2010. Land cover classification of remotely sensed satellite data using bayesian and hybrid classifier. International Journal of Computer Applications 7 (11), 1–4.

Prasad, A. M., Iverson, L. R., Liaw, A., 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems 9 (2), 181–199.

Pu, R., Gong, P., Tian, Y., Miao, X., Carruthers, R. I., Anderson, G. L., 2008. Using classification and NDVI differencing methods for monitoring sparse vegetation coverage: a case study of saltcedar in Nevada, USA. International Journal of Remote Sensing 29 (14), 3987–4011.

Pyzdek, T., 2003. The Six Sigma handbook: a complete guide for green belts, black belts, and managers at all levels. The McGraw-Hill, New York.

Qiu, F., Jensen, J. R., 2004. Opening the black box of neural networks for remote sensing image classification. International Journal of Remote Sensing 25 (9), 1749–1768.

Ramoelo, A., Skidmore, A., Cho, M., Schlerf, M., Mathieu, R., Heitkönig, I., oct 2012. Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne RapidEye sensor. International Journal of Applied Earth Observation and Geoinformation 19, 151–162.

RapidEye, 2012. Satellite Imagery Product Specifications. Tech. Rep. September.

Rapinel, S., Clément, B., Magnanon, S., Sellin, V., Hubert-Moy, L., 2014. Identification and mapping of natural vegetation on a coastal site using a Worldview-2 satellite image. Journal of Environmental Management 144, 236–246.

Reiche, J., Verbesselt, J., Hoekman, D., Herold, M., jan 2015. Fusing Landsat and SAR time series to detect deforestation in the tropics. Remote Sensing of Environment 156, 276–293.

Renza, D., Martinez, E., Arquero, A., 2013. A new approach to change detection in multispectral images by means of ERGAS index. IEEE Geoscience and Remote Sensing Letters 10 (1), 76–80.

Reynolds, L., 2006. Country pasture/forage resource profiles Malawi. Country profiles, Zimbabwe. Food and Agriculture Organisation, 1–28.

Ribeiro, N., Cumbana, M., Mamugy, F., Chaúque, A., 2012. Remote Sensing of Biomass in the Miombo Woodlands of Southern Africa: Opportunities and Limitations for Research. Remote Sensing of Biomass Principles and Applications, 77–98.

Ridd, M. K., Liu, J., 1998. A comparison of four algorithms for change detection in an urban environment. Remote Sensing of Environment 63 (2), 95–100.

Robertson, L. D., King, D. J., 2011. Comparison of pixel- and object-based classification in land cover change mapping. International Journal of Remote Sensing 32 (6), 1505–1529.

Rogan, J., Franklin, J., Roberts, D. A., 2002. A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. Remote Sensing of Environment 80 (1), 143–156.

Rosette, J. A., North, P. R. J., Suárez, J. C., Armston, J. D., sep 2009. A comparison of biophysical parameter retrieval for forestry using airborne and satellite LiDAR. International Journal of Remote Sensing 30 (19), 5229–5237.

Rossatto, D. R., Hoffmann, W. A., Franco, A. C., 2009. Differences in growth patterns between co-occurring forest and savanna trees affect the forest-savanna boundary. Functional Ecology 23 (4), 689–698.

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., Schott, J. R., Sheng, Y., Belward, A. S., Bindschadler, R., Cohen, W. B., Gao, F., Hipple, J. D., Hostert, P., Huntington, J., Justice, C. O., Kilic, A., Kovalskyy, V., Lee, Z. P., Lymburner, L., Masek, J. G., Mccorkel, J., Shuai, Y., Vogelmann, J., Wynne, R. H., Zhu, Z., 2014. Landsat-8 : Science

and product vision for terrestrial global change research. Remote Sensing of Environment 145, 154–172.

Ryan, C. M., Hill, T., Woollen, E., Ghee, C., Mitchard, E., Cassells, G., Grace, J., Woodhouse, I. H., Williams, M., jan 2012. Quantifying small-scale deforestation and forest degradation in African woodlands using radar imagery. Global Change Biology 18 (1), 243–257.

Ryan, C. M., Williams, M., Grace, J., 2010. Above- and belowground carbon stocks in a Miombo wooodland landscape of Mozambique. Biotropica 43, 423–432.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. a., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. Proceedings of the National Academy of Sciences 108 (24), 9899–9904.

Sadeghi, V., Ebadi, H., Ahmadi, F. F., may 2013. A new model for automatic normalization of multitemporal satellite images using Artificial Neural Network and mathematical methods. Applied Mathematical Modelling 37 (9), 6437–6445.

SAFNet, 2013. 9th Southern African Fire Network (SAFNet) Meeting. Tech. Rep. February, SAFNet, Morogoro.

Salford Systems, 2014. Data mining and predictive analytics.

Salvini, G., Herold, M., De Sy, V., Kissinger, G., Brockhaus, M., Skutsch, M., 2014. How countries link REDD+ interventions to drivers in their readiness plans: implications for monitoring systems. Environmental Research Letters 9 (7), 074004.

Sasaki, N., Putz, F. E., oct 2009. Critical need for new definitions of forest and forest degradation in global climate change agreements. Conservation Letters 2 (5), 226–232.

Scheba, A., Mustalahti, I., jan 2015. Rethinking expert' knowledge in community forest management in Tanzania. Forest Policy and Economics.

Schmidt, M., Lucas, R., Bunting, P., Verbesselt, J., Armston, J., mar 2015. Multi-resolution time series imagery for forest disturbance and regrowth monitoring in Queensland, Australia. Remote Sensing of Environment 158, 156–168.

Schmidt, M., Udelhoven, T., Gill, T., Röder, A., 2012. Long term data fusion for a dense time series analysis with MODIS and Landsat imagery in an Australian Savanna. Journal of Applied Remote Sensing 6 (1), 63512–63518.

Schneider, A., Friedl, M. A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environmental Research Letters 4 (4), 044003.

Schoene, D., Killmann, W., von Lupke, H., LoycheWilkiwe, M., 2007. Definitional issues related to reducing emissions from deforestation in developing countries.

Schoppmann, M. W., Tyler, W. A., mar 1996. Chernobyl revisited: Monitoring change with change vector analysis. Geocarto International 11 (1), 13–27.

Schwaiger, R., Huber, R., Mayer, H. a., 1995. Land Cover Classification of Landsat Images Using Problem Adapted Artificial Neural Networks. In: Binaghi, E., Brivio, P. A., Rampini, A. (Eds.), Soft Computing in Remote Sensing Data Analysis,. Milan.

Sebari, I., He, D. C., 2013. Automatic fuzzy object-based analysis of VHSR images for urban objects extraction. ISPRS Journal of Photogrammetry and Remote Sensing 79, 171–184.

Shalaby, A., Tateishi, R., 2007. Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt. Applied Geography 27 (1), 28–41.

Sheuyange, A., Oba, G., Weladji, R. B., 2005. Effects of anthropogenic fire history on savanna vegetation in north-eastern Namibia. Journal of Environmental Management 75 (3), 189–198.

Shimada, M., Ohtaki, T., 2010. Generating Large-Scale High-Quality SAR Mosaic Datasets: Application to PALSAR Data for Global Monitoring.

Shin, K.-S., Lee, Y.-J., 2002. A genetic algorithm application in bankruptcy prediction modeling. Expert Systems with Applications 23 (3), 321–328.

Shirima, D. D., Munishi, P. K. T., Lewis, S. L., Burgess, N. D., Marshall, A. R., Balmford, A., Swetnam, R. D., Zahabu, E. M., sep 2011. Carbon storage, structure and composition of miombo woodlands in Tanzania's Eastern Arc Mountains. African Journal of Ecology 49 (3), 332–342.

Shorrocks, B., 2007. The Biology of African Savannahs, 1st Edition. Oxford University Press, Oxford.

Shoshany, M., 2000. Satellite remote sensing of natural Mediterranean vegetation: a review within an ecological context.

Simonetti, E., Simonetti, D., Preatoni, D., 2014. Phenology-based land cover classification using Landsat 8 time series. Tech. rep., European Union, Luxembourg.

Singh, A., 1989. Review Article Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing 10 (6), 989–1003.

Singh, A., Harrison, A., jun 1985. Standardized principal components. International Journal of Remote Sensing 6 (6), 883–896.

SJP, 2009. Standard Joint Programme Document: UN-REDD Programme - Tanzania Quick Start Initiative. Tech. rep., UNDP, FAO, UNEP & Forestry and Beekeeping Division of the Minisrty of Natural Resources and Tourism.

Skidmore, A. K., Ferwerda, J. G., Mutanga, O., Van Wieren, S. E., Peel, M., Grant, R. C., Prins, H. H., Balcik, F. B., Venus, V., jan 2010. Forage quality of savannas Simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. Remote Sensing of Environment 114 (1), 64–72.

Skidmore, A. K., Turner, B. J., Brinkhof, W., Knowles, E., 1997. Performance of a Neural Network : Mapping Forests Using GIS and Remotely Sensed Data. Photogrammetric Engineering & Remote Sensing 63 (5), 501–514.

Slonecker, T. E., Johnson, B., McMahon, J., 2009. Automated imagery orthorectification pilot. Journal of Applied Remote Sensing 3 (1), 033552.

Soares-Filho, B. S., Nepstad, D. C., Curran, L. M., Cerqueira, G. C., Garcia, R. A., Ramos, C. A., Voll, E., McDonald, A., Lefebvre, P., Schlesinger, P., mar 2006. Modelling conservation in the Amazon basin. Nature 440 (7083), 520–523.

Sohl, T. L., 1999. Change Analysis in the United Arab Emirates: An Investigation of Techniques. Photogrammetric Engineering & Remote Sensing 65 (4), 475–484.

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., Macomber, S. A., 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? Remote Sensing of Environment 75 (2), 230–244.

Soudani, K., François, C., le Maire, G., Le Dantec, V., Dufrêne, E., may 2006. Comparative analysis of IKONOS, SPOT, and ETM+ data for leaf area index estimation in temperate coniferous and deciduous forest stands. Remote Sensing of Environment 102 (1-2), 161–175.

Sousa, C. H. R. D., Souza, C. G., Zanella, L., Carvalho, L. M. T. D., 2012. Analysis of RapidEye's Red Edge Band for Image Segmentation and Classification. In: 4th GEOBIA. Rio de Janeiro, pp. 518–523.

Souza, C. M., Roberts, D. a., Cochrane, M. a., oct 2005. Combining spectral and spatial information to map canopy damage from selective logging and forest fires. Remote Sensing of Environment 98 (2-3), 329–343.

Spinage, C., 2012. African Ecology: Benchmarks and Historical Perspectives. Springer, Berlin.

Stehman, S. V., 2009. Sampling designs for accuracy assessment of land cover. International Journal of Remote Sensing 30 (20), 5243–5272.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC bioinformatics 9, 307.

Suchenwirth, L., Förster, M., Cierjacks, a., Lang, F., Kleinschmit, B., 2012. Knowledge-based classification of remote sensing data for the estimation of below- and above-ground organic carbon stocks in riparian forests. Wetlands Ecology and Management 20 (2), 151–163.

Sun, G., Ranson, K. J., Guo, Z., Zhang, Z., Montesano, P., Kimes, D., 2011. Forest biomass mapping from lidar and radar synergies. Remote Sensing of Environment 115 (11), 2906–2916.

Sundström, R., 2010. Making the forest carbon commons: Tracing measures to reduce emissions from deforestation and forest degradation (REDD) in Angai Village Land Forest Reserve. Ph.D. thesis, University of Helsinki.

Sunseri, T. R., 2009. Wielding the ax: State forestry and social conflict in Tanzania, 18202000, 1st Edition. Ohio University Press, Athens.

Swetnam, R. D., Fisher, B., Mbilinyi, B. P., Munishi, P. K. T., Willcock, S., Ricketts, T., Mwakalila, S., Balmford, A., Burgess, N. D., Marshall, a. R., Lewis, S. L., mar 2011. Map-

ping socio-economic scenarios of land cover change: a GIS method to enable ecosystem service modelling. Journal of environmental management 92 (3), 563–74.

Tanzania Forest Conservation Group, 2007. South Nguru Mountains: a description of the biophysical landscape.

Tanzania Meteorological Agency, 2014. Tanzania meteorological agency - seasonal forecast.

Tanzania Natural Resources Forum, 2012. Working together for learning and action: shared experiences of the Tanzania REDD+ pilot projects. Tech. rep.

Taylor, J. C., Brewer, T. R., Bird, A. C., 2000. Monitoring landscape change in the National Parks of England and Wales using aerial photo interpretation and GIS. International Journal of Remote Sensing 21 (13-14), 2737–2752.

Tademir, K., Milenov, P., Tapsall, B., apr 2012. A hybrid method combining SOM-based clustering and object-based analysis for identifying land in good agricultural condition. Computers and Electronics in Agriculture 83, 92–101.

Terratec, 2012. Rapport for laser skanning. Tech. rep.

Thayn, J. B., 2012. Assessing vegetation cover on the date of satellite-derived start of spring. Remote Sensing Letters 3 (8), 721–728.

The Government of Tanzania, FAO, 2009. National Forestry Resources Monitoring and Assessment.

The Institute of Resource Assessment, 2010. The United Republic of Tanzania Resource manual on REDD basics. Tech. rep., Government of Tanzania, Dar es Salaam.

The Ministry of Natural Resources, 2012. Estimating Cost Elements of REDD+ in Tanzania. Tech. Rep. June.

The United Republic of Tanzania, 1997. National Environmental Policy. Tech. rep., Vice President's Office, Dar es Salaam.

The United Republic of Tanzania, 2010. National forestry resources monitoring and assessment of Tanzania (NAFORMA): Species List sorted by Latin names. No. December.

The United Republic of Tanzania, 2013. Tanzania in Figures 2012. Tech. rep., National Bureau of Statistics, Ministry of Finance, Dar es Salaam.

Themistocleous, K., Hadjimitsis, D. G., Retalis, A., Chrysoulakis, N., 2012. Development of a new image based atmospheric correction algorithm for aerosol optical thickness retrieval using the darkest pixel method.

Theodoridis, S., Koutroumbas, K., 2003. Pattern Recognition, 2nd Edition. Elsevier, San Diego.

Tomlinson, K. W., Poorter, L., Bongers, F., Borghetti, F., Jacobs, L., Van Langevelde, F., 2014. Relative growth rate variation of evergreen and deciduous savanna tree species is driven by different traits. Annals of Botany 114 (2), 315–324.

Tripathi, S. N., Dey, S., Chandel, A., Srivastava, S., Singh, R. P., Holben, B. N., 2005. Comparison of MODIS and AERONET derived aerosol optical depth over the Ganga Basin, India. Annales Geophysicae 23 (4), 1093–1101.

Tseng, M.-H., Chen, S.-J., Hwang, G.-H., Shen, M.-Y., 2008. A genetic algorithm rule-based approach for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 63 (2), 202–212.

Tso, B., Mather, P. M., 2009. Classification Methods for Remotely Sensed Data, 2nd Edition. CRC Press.

Turner, M. G., 2010. Disturbance and landscape dynamics in a changing world. Ecology 91 (10), 2833–2849.

Tzotsos, A., Argialas, S., 2006. A support vector machine approach for object-based image analysis. In: 1st International Conference on Object-based Image Analysis. Salzburg.

UN-REDD, 2011. The UN-REDD Programme strategy. Tech. rep., UN-REDD Programme Secretariat, Geneva.

UN-REDD Programme, 2015. UN-REDD Programme Strategic Framework 2016-20. Tech. rep., UN-REDD, Washington DC.

UNEP, 2001. Definitions; Indicative definitions taken from the Report of the ad hoc technical expert group on forest biological diversity.

UNEP, 2005. Facing the Facts: Assessing the Vulnerability of Africa's Water Resources to Environmental Change, early warn Edition. Nairobi.

UNFCCC, 2002. Report of the Conference of the Parties on its Seventh session, held at Marrakesh from 29th October to 10th November 2001. Tech. rep.

UNFCCC, 2010. Outcome of the work of the ad hoc working group on long-term cooperative action under the convention, draft decision. Tech. rep.

United Republic of Tanzania, 1997. Socio-economic profile: Lindi Region, volume 8 Edition. Socio-economic Profile. Planning Commission, Arusha Region.

United Republic of Tanzania, 2001. National Forest Programme In Tanzania 2001-2010. Tech. rep., Ministry of Natural Resources and Tourism, Dar es Salaam.

United Republic of Tanzania, 2002. United Republic of Tanzania. Country profile, Johannesburg Summit 2002. Tech. rep., United Nations, Johannesburg.

United Republic of Tanzania, dec 2003. Initial national communication under the United Nations Framework Convention on Climate Change (UNFCCC). Tech. Rep. 3, The United Republic of Tanzania.

United Republic of Tanzania, 2009. National framework for Reduced Emissions from Deforestation and Forest Degradation (REDD). Tech. rep., Vice President's Office, The Ministry of Natural Resources and Tourism, Dar es Salaam.

United Republic of Tanzania, 2012. REDD readiness progress fact sheet: Tanzania. Tech. rep.

United Republic of Tanzania, 2013a. National strategy for reduced emissions from deforestation and forest degradation. Tech. rep., Vice President's Office.

United Republic of Tanzania, 2013b. Piloting REDD in Tanzania. Tech. rep.

USGS, 2008. Global Land Cover Characterization.

USGS, 2014a. Landsat 5 History.
URL `http://landsat.usgs.gov/about_landsat5.php`

USGS, 2014b. Landsat and LDCM Headlines 2012.
URL `http://landsat.usgs.gov/mission_headlines2012.php`

USGS, 2015a. EarthExplorer.
URL `http://earthexplorer.usgs.gov`

USGS, 2015b. Landsat Missions: Imaging the Earth Since 1972.
URL `http://landsat.usgs.gov/about_mission_history.php`

Vapnik, V. N., 1999. An overview of statistical learning theory. IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council 10 (5), 988–999.

Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D., jan 2010. Detecting trend and seasonal changes in satellite image time series. Remote Sensing of Environment 114 (1), 106–115.

Verbesselt, J., Zeileis, A., Herold, M., 2012. Near real-time disturbance detection using satellite image time series. Remote Sensing of Environment 123, 98–108.

Vieira, M. A., Formaggio, A. R., Rennó, C. D., Atzberger, C., Aguiar, D. A., Mello, M. P., 2012. Object Based Image Analysis and Data Mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas. Remote Sensing of Environment 123, 553–562.

W Charitable Foundation, 2006. Tanzania: Lindi; Liwale. Tech. Rep. 1111441.

Walter, V., 2004. Object-based classification of remote sensing data for change detection. ISPRS Journal of Photogrammetry and Remote Sensing 58 (3-4), 225–238.

Wang, K., Franklin, S. E., Guo, X., He, Y., McDermid, G. J., 2009. Problems in remote sensing of landscapes and habitats. Progress in Physical Geography 33 (6), 747–768.

Wang, S., Zhou, Y., jul 2011. A study of 6S model used for atmospheric correction of MODIS image over Taihu Lake.

Warner, T. A., Nerry, F., 2009. Does single broadband or multispectral thermal data add information for classification of visible, near and shortwave infrared imagery of urban areas? International Journal of Remote Sensing 30 (9), 2155–2171.

Wasige, J. E., Groen, T. a., Smaling, E., Jetten, V., apr 2013. Monitoring basin-scale land cover changes in Kagera Basin of Lake Victoria using ancillary data and remote sensing. International Journal of Applied Earth Observation and Geoinformation 21, 32–42.

Waske, B., Benediktsson, J. A., 2007. Fusion of support vector machines for classification of multisensor data.

Waske, B., Van Der Linden, S., 2008. Classifying multilevel imagery from SAR and optical sensors by decision fusion.

Wasser, L., Day, R., Chasmer, L., Taylor, A., jan 2013. Influence of vegetation structure on lidar-derived canopy height and fractional cover in forested riparian buffers during leaf-off and leaf-on conditions. PloS one 8 (1), e54776.

Watson, R. T., Noble, I. R., Bolin, B., Ravindranath, N. H., Verardo, D. J., Dokken, D. J., 2000. Land use, land-use change and forestry: a special report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. Remote Sensing of Environment 117, 34–49.

Wieland, M., Pittore, M., 2014. Performance evaluation of machine learning algorithms for urban pattern recognition from multi-spectral satellite images. Remote Sensing 6 (4), 2912–2939.

Wijaya, A., Gloaguen, R., 2007. Comparison of multisource data support vector machine classification for mapping of forest cover.

Wijaya, A., Liesenberg, V., Gloaguen, R., 2010. Retrieval of forest attributes in complex successional forests of Central Indonesia: Modeling and estimation of bitemporal data. Forest Ecology and Management 259 (12), 2315–2326.

Wilson, E. H., Sader, S. a., 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery. Remote Sensing of Environment 80 (3), 385–396.

Winrock International, TerraCarbon, 2013. VCS Module VMD0007 REDD Methodological Module: Estimation of baseline carbon stock changes and greenhouse gas emissions from unplanned deforestation (BL-UP). Tech. Rep. May.

Wondrade, N., Dick, Ø. B., Tveite, H., 2014. GIS based mapping of land cover changes utilizing multi-temporal remotely sensed image data in Lake Hawassa Watershed, Ethiopia. Environmental Monitoring and Assessment 186 (3), 1765–1780.

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., Vogelmann, J., Wulder, M. A., Wynne, R., may 2008. Free access to Landsat imagery.

WRM, 2002. Tanzania: Biodiversity loss linked to IMF-promoted commercial agriculture and mining. WRM Bulletin (56).

Wulder, M. a., Masek, J. G., Cohen, W. B., Loveland, T. R., Woodcock, C. E., 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sensing of Environment 122, 2–10.

Wulder, M. a., Skakun, R. S., Kurz, W. a., White, J. C., 2004. Estimating time since forest harvest using segmented Landsat ETM+ imagery. Remote Sensing of Environment 93 (1-2), 179–187.

Wynne, R., 2006. Lidar remote sensing of forest resources at the scale of management. Photogrammetric engineering and remote sensing, 1310–1314.

Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. Journal of Plant Ecology 1 (1), 9–23.

Xu, E., Jia, Z., Wang, L., Hu, Y., Yang, J., 2014. Remote sensing image segmentation model based on the Otsu rule and K-means clustering algorithm. Information Technology Journal 13 (4), 690–696.

Xu, L. X. L., Zhang, S. Z. S., He, Z. H. Z., Guo, Y. G. Y., 2009. The comparative study of three methods of remote sensing image change detection. 2009 17th International Conference on Geoinformatics, 1595–1598.

Xu, M., Watanachaturaporn, P., Varshney, P. K., Arora, M. K., 2005. Decision tree regression for soft classification of remote sensing data. Remote Sensing of Environment 97 (3), 322–336.

Yan, G., Mas, J. F., Maathuis, B. H. P., Xiangmin, Z., Dijk, P. M. V., 2006. Comparison of pixelbased and object-oriented image classification approaches - a case study in a coal fire area, Wuda, Inner Mongolia, China. International Journal of Remote Sensing 27 (18), 4039–4055.

Yang, C., Everitt, J. H., Murden, D., 2011. Evaluating high resolution SPOT 5 satellite imagery for crop identification. Computers and Electronics in Agriculture 75 (2), 347–354.

Yang, X., Chen, L., Li, Y., Xi, W., Chen, L., 2015. Rule-based land use/land cover classification in coastal areas using seasonal remote sensing imagery: a case study from Lianyungang City, China. Environmental Monitoring and Assessment 187 (7), 449.

Yang, X., Ren, L., Singh, V. P., Liu, X., Yuan, F., Jiang, S., Yong, B., 2012. Impacts of land use and land cover changes on evapotranspiration and runoff at Shalamulun River watershed, China. Hydrology Research 43 (1/2), 23.

Yu, C., Yu, S., Shaohong, S., Shen, J., Huang, H. J., Yaohua, Y., 2010. An object-based change detection approach using high-resolution remote sensing image and GIS data.

Yuan, F., Sawaya, K. E., Loeffelholz, B. C., Bauer, M. E., 2005. Land cover classification and change analysis of the Twin Cities (Minnesota) metropolitan area by multitemporal Landsat remote sensing. Remote Sensing of Environment 98 (2-3), 317–328.

Zahabu, E., 2008. Sinks and sources: a strategy to involve forest communities in Tanzania in global climate policy. Ph.D. thesis, University of Twente.

Zahabu, E., Skutsch, M., 2008. The Likely Mechanism for Implementing REDD Policy in Tanzania. Think Global Act, 32–41.

Zhang, B., Chen, K., Zhou, Y., Xie, M., Zhang, H., 2010a. Research on Change Detection in Remote Sensing Images by using 2D-Fisher Criterion Function Method. In: ISPRS TC VII Symposium 100 Years. Vol. XXXVIII. pp. 697–702.

Zhang, B., Yao, Y., Cheng, W., Zhou, C., Lu, Z., Chen, X., Alshir, K., ErDowlet, I., Zhang, L., Shi, Q., 2002. Human-Induced Changes to Biodiversity and Alpine Pastureland in the Bayanbulak Region of the East Tienshan Mountains. Mountain Research and Development 22 (4), 383–389.

Zhang, L., Zou, B., Zhang, J., Zhang, Y., 2010b. Classification of polarimetric SAR image based on Support Vector Machine using Multiple-Component Scattering Model and texture features. Eurasip Journal on Advances in Signal Processing 2010.

Zhang, R., Zhu, D., 2011. Study of land cover classification based on knowledge rules using high-resolution remote sensing images. Expert Systems with Applications 38 (4), 3647–3652.

Zhang, Y., Xuemei, M., Liang, C., 2007. Hybrid change detection for watershed impervious surface using multi-time remotely sensed data.

Zheng, D., Wallin, D., Hao, Z., 1997. Rates and patterns of landscape change between 1972 and 1988 in the Changbai Mountain area of China and North Korea. Landscape Ecology 12 (541), 241–254.

Zhou, Q., Li, B., Chen, Y., 2011. Remote sensing change detection and process analysis of long-term land use change and human impacts. Ambio 40 (7), 807–818.

Zhou, W., Huang, G., Troy, A., Cadenasso, M. L., 2009. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. Remote Sensing of Environment 113 (8), 1769–1777.

Zhu, Z., Wang, S., Woodcock, C. E., mar 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 47, 8, and Sentinel 2 images. Remote Sensing of Environment 159, 269–277.

Zhu, Z., Woodcock, C. E., mar 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment 118, 83–94.